R
724
U561G
1963
v.3

ring
ss to
th Care

The Ethical Implications of Differences in the Availability of Health Services

Volume Three: Appendices Empirical, Legal, and Conceptual Studies

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research

Library of Congress card number 83-600501

For sale by the Superintendent of Documents
U.S. Government Printing Office
Washington, D.C. 20402

Securing Access to Health Care

The Ethical Implications of Differences in the Availability of Health Services

Volume Three: Appendices
Empirical, Legal, and Conceptual Studies

March 1983

United States

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research

President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research

Morris B. Abram, M.A., J.D., LL.D., *Chairman,*
New York, N.Y.

H. Thomas Ballantine, M.D., M.S., D.Sc.*
Harvard Medical School

George R. Dunlop, M.D.
University of Massachusetts

Bruce K. Jacobson, M.D.
Southwestern Medical School

John J. Moran, B.S.
Houston, Texas

Arno G. Motulsky, M.D.
University of Washington

Daher B. Rahi, D.O.
St. Clair Shores, Michigan

Seymour Siegel, D.H.L.
Jewish Theological Seminary of America, New York

Lynda Smith, B.S.
Colorado Springs, Colorado

Kay Toma, M.D.
Bell, California

Charles J. Walker, M.D.
Nashville, Tennessee

* A dissenting statement by Commissioner Ballantine appears at the end of Volume One: Report.

Staff

Alexander M. Capron, LL.B., *Executive Director*

Deputy Director
Barbara Mishkin, M.A., J.D.

Assistant Directors
Joanne Lynn, M.D., M.A.
Alan Meisel, J.D.

Professional Staff
Mary Ann Baily, Ph.D.
Dan Brock, Ph.D.
Allen Buchanan, Ph.D.
Andrew Burness, M.B.A.
Kathryn Kelly, M.S.W.
Susan Morgan
Marian Osterweis, Ph.D.
Renie Schapiro, M.P.H.
Daniel Wikler, Ph.D.

Research Assistants
Katherine Locke
Jeffrey Stryker

Consultants
Bradford H. Gray, Ph.D.
Dorothy Vawter

Administrative Officer
Anne Wilburn

Editors
Sharin Sachs
Linda Starke

Support Staff
Florence Chertok
Gretchen Erhardt
Ruth Morris
Clara Pittman
Kevin Powers
Nancy Watson

President's Commission Commonwealth Fellows and Student Interns
Susan Formaker (1982)
Kenneth Kim (1982)
Eddie Lockard (1982)
Stephen Massey (1982)
Lisa Rovin (1982)
Mindy Werner (1982)

This volume consists of materials prepared as background for the Commission's report on this subject.

Table of Contents

Volume Three: Appendices (Empirical, Legal, and Conceptual Studies)

Part III: Ethical Implications of Health Care Distribution 1

M. Equity, Access, and the Costs of Health Services (Bruce C. Vladeck) 3
Definitions 4
Past Accomplishments 6
The Erosion of Access 8
Issues 13
Conclusions 16

N. Equity of Access to Medical Care: A Conceptual and Empirical Overview (Lu Ann Aday and Ronald M. Andersen) 19
Introduction 19
Conceptual and Methodological Issues in Equity of Access 21
Trends in the Profile of Equity 29
Health Policy and the Current Profile of Equity 49

O. Uninsured and Underserved: Inequities in Health Care in the U.S. (Karen Davis and Diane Rowland) 55
Who Are the Uninsured? 56
Utilization of Health Services by the Uninsured 62
Policy Implications 73
Recommendations 75

P. The Medicaid Program in Transition (Gerald R. Connor) 77
Background 78
The Medicaid Program 80
State Medicaid Programs 88
1981 Medicaid Amendments 94
Possible State Actions Under 1981 Amendments 98
1983: The Debate Continues 101

Q. Access to Private Physicians for Public Patients: Participation in Medicaid and Medicare (Janet B. Mitchell and Jerry Cromwell) 105
Medicaid Participation 106
Medicare Assignment 120
Summary and Policy Implications 125
Recommendations 128

R. Equity in the Distribution of Quality of Care (Leon Wyszewianski and Avedis Donabedian) 131
Framework and Definitions 133
The Empirical Evidence 143
Summary and Discussion 160

S. Health Status as a Measure of Need for Medical Care: A Critique (John Yergan *et al.*) 173
Health Status Measurement 174
The Link Between Health Status and Access 177
Measures of Access Based on Health Status 179
Recommendations for a Comprehensive Measure 184

T. Concepts of Medical Underservice: A Review and Critique (Eugenia S. Carpenter) 189
Introduction 189
Components of Shortage or Underservice Measures 190
Construction of Comprehensive Measures of Shortage or Underservice 198
State Approaches to Defining Medical Underservice 205
Appropriateness of Existing and Proposed Indices of Underservice 211
Physician Location Decisions as an Element of Medical Underservice 215
Conclusions 220

Part IV: Allocation in Varied Medical Settings 223

U. The Doctor, The Hospital, and the Definition of Proper Medical Practice (Anthony L. Komaroff) 225
The Doctor and the Hospital Administration 225
Benefits, Risks, and Costs 227
Proper Medical Practice and a Doctor's Decisions 232
Variability and Misutilization in Medical Practice 236
Interventions 244
Involving the Doctor in Hospital Cost Control 247
Conclusion 251

V. The Rationing of Hospital Care (E. Richard Brown) 253
The Changing Allocation of Hospital Care 255
Cost Containment and the Rationing of Hospital Care 272

W. The Allocation of Resources for Medical Intensive Care (Albert G. Mulley) 285
History, Availability, and Costs 286
ICU Patients 292
Rationing 299
Equity, Ethics, and Policy Implications 305

X. Health Maintenance Organizations and the Rationing of Medical Care (Harold S. Luft) 313
Economic Incentives in Various Medical Settings 314
Socioeconomic Differences in the Effects of Alternative Allocation Systems 324
Potential Effects Under Increased Competition 331
Ethical and Policy Implications of Rationing of Medical Care in HMOs 335

Part V: Legal Implications of Allocation Policies 347

Y. Rationing Medical Resources: A Constitutional, Legal, and Policy Analysis (James F. Blumstein) 349
Introduction 349
Defining Problems and Identifying Issues 350
Constitutional Constraints on Rationing Medical Care 358
Legal Implications of Private Rationing Decisions 388
Conclusion 394

Z. Rationing "Normal" Health Care: The Hidden Legal Issues (Rand E. Rosenblatt) 395
Rationing Health Care: Images and Realities 397
Professor Blumstein and the Problem of the Vanishing Poor 400
The Failure of Malpractice Law to Protect the Poor 404
Rationing, Excessive Health Services, and Malpractice Law 408
Conclusion 411

AA. Malpractice Liability and the Rationing of Care (Peter H. Schuck) 413
Altering Substantive Norms 415
Changing the Nature of the Provider 416
Expanding Financial Incentives 416

Volume One: Report and **Volume Two: Appendices (Sociocultural and Philosophical Studies)** are available from the Superintendent of Documents.

Ethical Implications of Health Care Distribution

III

Equity, Access, and the Costs of Health Services

Bruce C. Vladeck, Ph.D.*

Sometime in the last decade, the axis of health policy debate in the United States underwent a fundamental shift. Expanded availability of health care services to the population as a whole or identified groups within it had been the major issue in health policy debate since at least the early days of the New Deal. That issue found expression not only in the 30-year struggle that culminated in the enactment of Medicare and Medicaid, but also in governmental initiatives to expand the supply of hospitals, other health facilities, and then health personnel. Contemporary debate, on the other hand, often takes as a given the oversupply of at least some medical services, tends to focus primarily on issues of cost and cost containment, and frequently involves the proposition that Americans, or at least some of them, use too many health services.

To a considerable degree, Medicare and Medicaid achieved the primary objectives of increased availability of services to beneficiary populations; the recent shift in policy focus might be depicted as an inevitable outcome of success at addressing earlier problems. In an altenative formulation, one could take the frequently heard argument that current problems of cost and expenditure control are the direct results of those programs. As in other areas of social policy, this argument goes, we have fallen victim to the unintended consequences of noble objectives.

* Assistant Commissioner, Health Planning and Resources Development, New Jersey State Department of Health, Trenton, New Jersey.

This study, which was originally prepared for the Commission, is reprinted with permission from MEDICAL CARE, Vol. 19, No. 12 (Supplement), 1981.

February 1981.

This latter argument is explored in more detail later, but it is important to emphasize at this point that policy debate, perhaps especially in health care, often assumes a certain pendulum-like character. The cycles of oscillation between concerns for cost and access may be relatively lengthy, but the pendulum does appear to be swinging. Before it swings too far, it is not only appropriate but necessary to explore what, in fact, has happened to the availability of health services, why availability has been replaced by cost as the focus of policy concern, and whether the often-assumed diametrical relationship between the two is really grounded in sound evidence or logic.

The basic contention of this paper is that the actual improvements in access to health care that occurred over the last 15 years have already begun to erode alarmingly. Further, those improvements still left a situation considerably short, in many important respects, of what might be considered desirable. This erosion in availability arises to a considerable degree from efforts at cost containment; yet cost containment, properly conceived, need not be inimical to improved service availability. The short-run outlook appears to hold reduced availability as a product of cost containment efforts, which, it can be argued, is a function of political choices that are far from unavoidable.

To make this argument, this paper first very briefly summarizes changes in patterns of medical care utilization since the inception of Medicare and Medicaid; somewhat less briefly surveys current problems in the availability of health services and makes some observations about the implications of some current policy trends; and then turns to a somewhat more abstract exploration of three critical issues. These might be posed as: the distinction between equality of access and the assurance of access to minimally "needed" services; the gap between equity in volume of services and equity in quality of services; and the nexus between access/availability and cost containment. The implications of these issues and their discussion provide the basis for some concluding statements.

Definitions

First, the thicket of definitions must be entered. Terms such as "access," "availability," and "equity" have themselves been the subjects of prolonged and often complex discussions. There is no need, however, for the narrative purposes of this paper, to arrive at definitive and unalterable generic definitions. On the other hand, it is necessary for the reader to have some notion of what the writer is talking about, so the following, rather commonsensical usages are observed.

Access/Accessibility. These terms describe a relationship between a potential user (consumer) of a service and the service, such that a potential user who desires or "needs" the

service is able to receive it without being barred for reasons of social status, income, ability to pay, place of residence, or other factors believed extraneous to the appropriate delivery of health care services. In this regard, access and accessibility are often defined in the negative, by the absence of obstructions that may prevent users from receiving services. For the purposes of this paper, no presumption is made about the "objective" need for or desirability of services, although services deemed unnecessary or undesirable by service providers do not, in this definition, constitute problems of access.

Availability. In contrast to accessibility, availability describes the characteristics of a service rather than of the relationship between services and their potential consumers. It is thus closely analogous to the economic notion of "supply," although in discussions of health services the question of "availability to whom?" is often of central concern; while in classic economics "supply to whom?" is almost literally meaningless. Although access is often defined in the negative, "availability" implies a positive. One can talk about a service for which there are no barriers to access, but which is still unavailable because there is no supply. Conversely, the extent to which services that are available are used provides at least a partial measure of their accessibility.

Cost Containment. Two quite different uses of this term are commonly encountered, and it is critical to distinguish between them. At the most basic level, cost containment constitutes a reduction in the rate of increase (because, in the health sector, one apparently can never speak of cost *reductions*) in the amount of resources employed to produce a given volume of services. More frequently, however, cost containment refers to a reduction in outlays by some particular payor or payors (most often governmental), which is often accomplished by a reduction in services. This sort of ersatz cost containment often produces increased expenditures for other parties; often it is achieved even while the unit costs of services continue to increase.

Equity. Anyone but a courageous philosopher hesitates to define "equity" simply, but for the purposes at hand, one can fall back on the Aristotelian principle that like cases should be treated alike and unlike cases unlike. In the provision of health services, equity is an attribute of a system that provides roughly similar services to those with similar health problems, and appropriately dissimilar services to those with dissimilar problems. The question of a system's equity is, of course, begged in the absence of agreement as to what similar or dissimilar problems, or what "appropriately" dissimilar services, may be.

Need. Following Boulding, need is that which professionals determine merits their intervention.[1] To avoid the problem of self-serving individual providers, one could borrow lan-

guage from malpractice law and define need as that which the prevailing community standard of best practice determines merits professional intervention.

Past Accomplishments

Historically in this country and elsewhere, two fundamental relationships described the connection between socioeconomic status and health services. First, income (and related measures such as education) is strongly related to health. At the same time, income is closely correlated with the use of health services. The poorer the individual, the more likely he is to have health problems, yet the fewer health services he is likely to receive.

Medicare and Medicaid changed the latter of those equations, and did so with astonishing rapidly given the magnitude of the social transformation. In 1964, middle-class people visited physicians 20% more often than the poor; by 1975, the poor had 20% more visits. More striking still are the data for children. Physician visits by poor children increased by 70% between 1964 and 1975, attaining parity with middle-class children in the latter years. Although the elderly had always used health services more intensively than younger people, hospital use rates skyrocketed after the inception of Medicare, and income differentials in service use fell as dramatically as among the non-elderly.[2]

Poor Americans now use roughly the same amount of physicians' services, and more hospital and nursing home services, than the nonpoor. In general, use of health services across income classes is substantially more equal than it was two decades ago. Yet controlling for measures of "need" or degrees of illness, the correlation between poverty and lesser use remains, although it has also shrunk. That is to say, a system that provides the poor with the same volume of services as the nonpoor has not attained equity, by the definition above, so long as poverty remains a powerful predictor of ill health, as it still does.[3]

Moreover, there have always been substantial gaps in the impact of Medicare and Medicaid on their target populations, exceptions to the general pattern of increasing equality, if not equity. Davis and Schoen have documented well the continuing shortfalls in service use by the rural poor and elderly and the Southern poor.[2] Medicaid eligibility requirements and coverage patterns continue to vary enormously among the states. And Medicaid, like other income-related social welfare programs, has created serious "notch" problems; an individual who barely fails to meet eligibility requirements receives no benefits but an otherwise identical individual who does meet those requirements may enjoy an extensive array of services. The largest

group of those affected by such "notch" problems, and the one that raises the most compelling issue of public policy, are poor children in two-parent households in the majority of states where only poor children in one-parent households are eligible for Medicaid.

Indeed, the availability of Medicaid for the poor has, in a perverse sort of way, exacerbated the access problems of those who are not poor (by a state's Medicaid eligibility definitions) but nearly so. Twenty-five million Americans have no health insurance; a large proportion of that number are the near-poor, and the uninsured are disproportionately (compared with the general population) rural and disproportionately black. The uninsured are younger than the rest of the population, but not apparently healthier.[4] Although specific data linking income, insurance status, and service use are not readily available, the poor now use health services of all kinds more than the near-poor. Indeed, the curve plotting income versus service use describes a rather wide and somewhat lopsided "U," with the highest use among the poorest and the wealthiest, and the lowest use among the almost poor.[2]

Notwithstanding these concerns, Medicare and Medicaid have contributed to an enormous shift in relative patterns of service use. At the same time, however, they have also contributed to the enormous inflation in health care costs that now bedevils policymakers. Some of this cost increase was anticipated; certainly, those who worked for their enactment expected that Medicare and Medicaid would engender increased utilization of services by their beneficiaries—which was what they were intended to do—and some argued that such an increase in demand, in the face of a relatively fixed supply of facilities and professionals, would have an additional inflationary impact. No one foresaw, at least not in writing, how severe the inflation would be.

The extent to which Medicare and Medicaid have independently contributed to cost inflation in health services, apart from a variety of other inflationary causes, is a subject of considerable debate. Hospital costs began increasing at a rate higher than overall inflation *before* 1965, and have continued to grow far faster than the increase in Medicare utilization.[5] Yet peculiarities of Medicare and Medicaid reimbursement policies have undoubtedly had an independent inflationary effect.[6] The supply of physicians has increased dramatically over the last decade as an outgrowth of policy decisions largely independent of Titles 18 and 19; because the growth of physician supply has outpaced the growth in service use, fees have risen according to the rather peculiar, supply-driven economics of medical care. Medicare and Medicaid have ridden on the tail of the cost-inflation tiger, although they have undoubtedly given it an occasional pull or two as well.

It is clear that most efforts to contain the program costs arising from the compounded effects of increased utilization and increased prices have involved service reductions or the imposition (or reimposition) of other barriers to access rather than direct attempts to control prices or increase efficiency. That has been the case with Medicaid in particular, where cutbacks in eligibility and service coverage began as early as 1967, and where "cost containment" has almost always referred to *program* cost containment.[7] Whether or not Medicare and Medicaid are primary causes of the inflation in health care costs, inflation has been a primary cause of limitations on their impact.

The Erosion of Access

Largely because of cost inflation, the improvements in access achieved by Medicare and Medicaid by the mid-1970s may have represented the high-water marks of such improvements, and the tide may have begun to recede even before last year's election. Medicaid enrollment fell between 1975 and 1979, despite general economic stagnation and increases in the population. Certainly, in most states, income eligibility standards for Medicaid have not increased nearly as fast as inflation. The most dramatic case is that of New York, with the nation's largest Medicaid program, where eligibility standards were essentially not changed between 1974 and 1980.*

As the population ages, Medicare enrollment continues to increase, but the proportion of all health expenditures for the elderly paid by it has fallen continually—to roughly 40% in 1978 (although another 20% of health care costs for the elderly is borne by other public programs, primarily Medicaid).[8] That phenomenon arises primarily from the falling proportion of physicians who accept Medicare assignment (payment in full) and the increasing proportion of long-term care costs, especially for nursing homes, in health expenditures for those over 65.

These aggregate figures conceal more than they reveal about problems in the accessibility and availability of health services, however; to get a fuller picture of those problems, a more detailed, descriptive approach is necessary, one that separately records the variety of persistent and new barriers to the receipt of services.

The Urban Poor. Although Medicaid led to a substantial increase in the quality of services received by the poor, it had a less dramatic effect on the character of the services delivered—

* For an extended discussion of this topic, see BC Vladeck, W Carr, *Health Policy.* In: SETTING MUNICIPAL PRIORITIES, 1982, R Horton, C Breacher, eds. (New York: Basic Books).

the "two-class" system of urban medical care has never been entirely eliminated. More recent program cutbacks have begun to lead to a return to past patterns of differential kinds of services for the poor and near-poor. Historically, the urban poor have been much more reliant on hospital-based or freestanding clinics—as opposed to private practitioners—for outpatient services, and public rather than private hospitals for inpatient services. In the early years of Medicaid, these patterns were altered, and use by those eligible for Medicaid of both private physician visits and private hospitals increased dramatically. But physician fee schedules in most of the larger Medicaid programs have been increased only slowly, if at all, over the last decade, and private physicians—like pharmacists, banks, and supermarkets—have been part of the general migration of services from inner-city areas. In many parts of the country, therefore, the urban poor are once again reliant on clinic services.

Apart from whatever esthetic differences such a pattern might imply, the fact remains that the poor—and to an even greater extent, the near-poor, who lack even the limited access to private physicians Medicaid provides—have no continuing relationship with individual physicians. This means that access to other services resides in the hands of bureaucratic strangers. The poor and near-poor are more likely to be admitted to hospitals through emergency room than a scheduled admission, and more likely to be sicker when they are admitted. They are more likely to be treated by a foreign medical graduate, or a house staff physician still in training (often both) than a fully trained American graduate. Their drug prescriptions are written by physicians with less personal knowledge and understanding of the patients' characteristics and problems, and they are likely to be seen on a follow-up visit by a physician other than the one who treated them originally.

Moreover, particular institutions that the urban poor have traditionally relied on for care, which were "their" hospitals or clinics, have begun to disappear, creating problems of access even when the total volume of services in the community falls no faster than the population. Public hospitals have closed in Philadelphia, St. Louis, and New York, and inner-city private hospitals throughout the country have closed or relocated. The accelerating shrinkage of the inner-city population has meant that inpatient capacity has fallen less than inpatient demand in most cities, so the acutely ill poor person who appears in an emergency room is still likely to be admitted and treated, even if in a strange hospital. But the availability of non-emergency outpatient services in many cities is much more tenuous; if nothing else, longstanding patterns of care and service utilization are irrevocably disruptive by the closure of a Philadelphia General or Homer Phillips Hospital, or the relocation of a Wilmington Medical Center.

The near-poor who are not eligible for Medicaid never have achieved equality in service usage with the more affluent, and their number appears to have swelled with the economic problems of the last few years, the slow growth in Medicaid rolls, and in a number of cities, it is widely reported, the dramatic increase in the number of undocumented aliens. There is no systematic data on the use of health services by these people, but both anecdotal reports and common sense suggest that many of them indeed encounter very severe access problems.

Problems in access to services may also increase as a result of the dramatic southwestward shift in the nation's population. Although the poor and near-poor have not joined the migration to the extent the rest of the population has, the population has been shifting from states with relatively generous Medicaid programs to states with relatively less generous programs. The pattern of reliance by the poor and near-poor on a small number of chronically underfunded public institutions has always been strongest in the South and Southwest, and Medicaid does not seem to have altered it significantly.[9] Further, the larger, and growing, market share of for-profit hospitals in the states with the greatest population growth portends access problems if the reputed reluctance of such institutions to provide charity care persists.

The poor, both urban and rural, are also more likely to have the types of health problems where availability problems have persisted, notwithstanding the progress of the last two decades. Rehabilitation services of many kinds are still undersupplied in much of the nation, and the poor and near-poor are more likely to have disabling chronic diseases. More dramatically, the reduction in state mental hospital populations has not been matched, in most of the country, by a concomitant increase in community-based services, and there are significant shortages of mental health personnel everywhere but the most affluent communities. If it was once true that the rich went to private sanitariums, the upper-middle class to psychiatrists, and the poor to state hospitals, the poor are still being admitted to state hospitals in large numbers but do not stay very long and are quickly discharged back to underserved communities.

It can hardly be said that policymakers have been unaware of continuing difficulties in access or availability for the inner-city poor and near-poor, nor that they have not sought to do something about them. There have been major Federal investments in community health centers under a variety of programmatic names, in Area Health Education Centers, in The National Health Services Corps, and a variety of other programs. Although falling short of their own definitions of the need for services, these programs have filled critical gaps in service availability in many urban areas from which private providers have departed, and have, in many cases, offered low-income

residents services qualitatively equal, if not superior, to those provided the middle-class in the "private sector." These services were, however, major casualties of the Federal budget reductions in 1981, and service levels will undoubtedly be reduced in the coming years.

In short, a low-income resident of an inner city is likely to have access to relatively high-quality emergency services in acute life-threatening episodes, and some form of access to episodic acute services when seriously ill. The availability of routine medical, dental, and psychiatric services, and their benefits—both for the individual and the wider community—is more likely to be a matter of geographic happenstance. If the individual resides near a well-managed health center, or a still-solvent community hospital that has undertaken to upgrade its outpatient services, or in a state whose Medicaid program still encourages participation by private physicians, access to health services comparable to those received by the bulk of the population may be relatively easy. Among the nation's urban poor, individuals so situated are likely to be in the minority.

The Rural Poor. In our society, those most likely to encounter access difficulties in the literal, ordinary use of the term are the rural poor, especially the minority rural poor. The number of counties without a physician in practice has increased since 1965, and although rural people see physicians almost as often as urban dwellers, they are far more likely to have to travel considerable distances to receive medical care, and wait a long time for service once they arrive.[10]

The publicly available data often do not permit such fine distinctions, but there is good reason to believe that travel distance, time costs, and the remnants of segregation make access most difficult for the rural minority poor, and rising fuel costs can only exacerbate their problems. Again, those most acutely and treatably ill are likely to have fewer problems of access, but those with chronic diseases or in need of preventive services are probably less well served than their urban counterparts.

As with urban health services, there has been a substantial Federal investment over the last decade to make services more available in rural areas, and to the rural poor in particular. Especially notable are those, such as the treatment of Rural Health Clinics under 1978 Amendments to Federal Medicare and Medicaid law, that seek to expand the availability of services to rural residents through reliance on nonphysician primary care personnel: nurse practitioners and physician's assistants. This strategy raises interesting ethical questions because, although all the evidence suggests that such paraphysician providers render services of equivalent quality to those rendered by physicians, they are still not physicians.

A different kind of policy problem, but one perhaps no less excruciating, is raised by the increasing economic plight of many small rural hospitals. Under current circumstances, it becomes harder every year for such institutions to operate without financial losses. Perhaps more importantly, it is increasingly difficult for institutions of such a size and character to provide the range of technologic services on which contemporary medical practice is ever more reliant, and which patients have a right to expect. On the other hand, in many rural counties a small hospital is the only health care resource of any sort.

The policy dilemma is this: continued subsidization of such facilities is expensive, and may leave at least some rural residents with lower quality (or at least lower technology) services than they could receive at a distance no greater than they are accustomed to travel for a shopping expedition or to go to the movies. But the closure of such facilities can leave relatively stable, homogenous communities without a major community resource to call their own, without the ability to attract new health care professionals, and without easy access to services adequate to meet at least some of their needs. Nor does it appear that technology will provide the fix for this dilemma; hopes that two-way telecommunications or time-shared computer systems would be the salvation of rural hospitals now appear realistic for only larger rural institutions—although the growth of managerial systems based in large urban hospitals, such as the Sisters of Mercy Corporation or Inter Mountain Health Care, may provide more realistic grounds for optimism.

The Elderly. The elderly, as a group, have been the principal beneficiaries of public policy in health care (and other sectors) over the last two decades, but serious problems of access and availability persist, especially for services not covered by Medicare. For the poor elderly, prescription drugs, eyeglasses, and dental services are much more difficult to come by than they are for the near-poor, and here cost is clearly the major barrier.[11]

The most dramatic and pervasive problems involve long-term care. Here the problem is twofold, although the components are closely interrelated. In most communities, there is an inadequate supply of noninstitutional long-term care services for all but the most affluent. Secondly, the poor elderly have substantially greater difficulty obtaining access to institutional services. In almost every state, Medicaid recipients go to the end of the nursing home queue, regardless of their need for services, whenever there is any unfulfilled private demand.[12] At the same time, private demand is undoubtedly "excessive" because of the unavailability of alternative services.[13] There is also a widespread belief that, for individuals seeking nursing home care, access is inversely related to need. That is, nursing homes prefer to admit the less disabled, less expensive patients.[14]

It should also be emphasized that patterns of use by the elderly, even for Medicare recipients, vary dramatically by race. Nonwhite Medicare enrollees are 15% less likely to be admitted to hospitals (although their stays, once they are admitted, are longer), 15% less likely to see physicians, and fully 30% less likely to be admitted to nursing homes for covered services.[10] These differentials are partially explained by the higher proportion of nonwhite elderly residing in rural areas, especially in the South, and by the continuing linkage of poverty and race. They are not explained by health status, because the nonwhite elderly, like nonwhites in general, experience more illness and disability than white persons of similar ages.

Issues

This abbreviated summary of remaining access problems hardly does justice to the literally thousands of instances, occurring every day in every part of the nation, in which people who need and want medical attention fail to receive it because of economic barriers, transportation difficulties, continuing patterns of racial or ethnic discrimination, or the unavailability of services. It should provide, however, some background for a consideration of the troubling policy questions surrounding the persistence of access problems.

The first of these might be described as the distinction between equity of access and the ability to receive services when they are needed. If, as is often contended, 80% of visits to physicians involve either self-limiting conditions, untreatable problems,or the "worried well," then it may be highly misleading to evaluate health policy in terms of the relative numbers of visits by income or social status. If some proportion of all surgical procedures are of marginal necessity, or even entirely unnecessary, then the 133% increase in surgical rates for the poor between 1963 and 1970 may represent a very mixed blessing.[2] It has been argued frequently that the least serious access problems are encountered by those who are most acutely ill; in a time of increasingly scarce resources, that might suggest that access as a policy problem should receive short shrift indeed.

If the problem of health care costs is perceived as one of the overutilization of services (such as hospital days and physician visits) then it may be rational, in benefit-cost terms, to attempt to *restrict* access to noncritical services by the beneficiaries of public programs (or the population as a whole) through the use of devices such as increased copayments, mandatory prior authorization of services, or rationing on the basis of waiting times or geographic inaccessibility.

There are three fundamental problems with that logic. First, it has never been adequately demonstrated that such methods restricting access, with the exception of prior authorization for some services, adequately distinguish between ben-

eficial and nonbeneficial utilization. That is to say that neither the consumer nor the bureaucrat is generally capable of equating the marginal cost of a service with its marginal benefit. High copayments may discourage physician visits by the worried well, but they may also discourage some of the fraction of visits that detect serious diseases at their earliest and most treatable stages. From a benefit-cost perspective, ten physician visits for self-limiting upper respiratory infections may be a cheap price to pay for one visit that prevents a severe pneumonia.

Second, increased utilization explains only a small fraction of the increase in health care costs over the last two decades—perhaps as little as 10% of the increase in hospital costs, which consume 40% of all health expenditures.[5] Indeed, such are the economics of health services that declining utilization is more likely, at least in the short run, to lead to increased unit prices rather than reduced total costs.

Most basically, serious ethical issues are raised by policy mechanisms that, on the one hand, seek to insure access to "needed" care for those unable to pay for it themselves while, on the other hand, denying them the opportunity to participate in the health-care market in the same way that other citizens do. The guarantees of free choice of providers contained in Medicare and Medicaid, which are now under increasing attack by those preoccupied with budgetary problems, represent more than symbolic rhetoric or a compromise gesture to the provider community. They embody a very profound and deeply held commitment to the belief that, if health care is a merit good that should be available to all regardless of economic or social circumstance, then it should be available in equal forms under equal conditions.

To be sure, those responsible for expanding taxpayers' funds have a special responsibility to insure that those funds are not spent wastefully or on unproductive services. But if, under current circumstances, much health care consumption is unavoidably very wasteful, there is a serious equity problem in only barring wasteful consumption to beneficiaries of public programs.

A second issue is raised by the qualitative differences in services received by the poor and minorities as opposed to the white nonpoor. Medicare and Medicaid made a considerable dent in the "two-class" system of health services, but they did not eliminate the class system altogether. This especially thorny issue has at least two components. To begin with, poor black and Hispanic people live in poorer communities with poorer institutions and fewer resources of all kinds. It may be unrealistic to expect health services to be so different, in their qualitative aspects, from schools, parks, or libraries. An unrealistic burden also may be imposed on health financing programs to expect them to overcome the whole set of interlocking social,

economic, and demographic forces contributing to the resource deficits of poor communities.

Further, although the poor may be more likely to receive services in underfunded public facilities, in clinics, or from paraphysicians, there is no good evidence that, in the narrow sense of medical care quality as opposed to the broader, more subjective sense, such services are, or always are, inferior. Certainly, the brittle diabetic treated in a speciality clinic of a major academic medical center may be receiving a more sophisticated service than a similar patient treated in the office of an overworked general practitioner who has not kept current with developments in clinical practice, although the latter patient may receive care closer to home, after a shorter wait, in a much friendlier environment.

On the other hand, many people who view socioeconomic differences in the quality of many services received as an inevitability recoil at the persistence of such differences in the health sector, and they may have a very strong moral case. Health care *is* different; at least some small fraction of the time, it is literally concerned with life and death.

Finally, there is the issue of cost containment. There is no question but that, to date, cost containment narrowly defined has had a disproportionately adverse effect on the poor and members of minority groups. When hospitals are closed as a result of regulatory or reimbursement pressures, they tend to be inner-city or rural hospitals; when reimbursement rates are regulated, they tend to be the rates paid by public programs; certificate of need controls may discourage supply increases in inner cities as stringently as in the suburbs.

Again, this problem is at least partially one of definition. True cost containment must affect those who provide services to the entire population, or it will not work; programmatic cost containment adversely affects the poor and the elderly because they are the principal beneficiaries of public programs.

The astonishing increases in program costs for Medicare and Medicaid over the last decade have arisen, not from increased utilization by program beneficiaries, but from increased unit prices paid to providers. Inflation in hospital costs, nursing home rates, and physicians' fees continues to outpace the general rate of inflation, and accounts for the "uncontrollability" of public health insurance programs. Because total program expenditures can be conceived of as a product of the number of service units times the unit price of services, it is clear that cost containment must come either through reductions in service levels or control of price inflation.

Reduction of services in Medicare and Medicaid will inevitably mean abandonment of much of the gain in access to health services by the poor and elderly that has occurred over

the past 15 years. Reduction in reimbursement levels by public programs alone will also exacerbate access problems, as providers will choose to serve patients whose insurance or own pockets can pay higher rates. Controlling the prices charged to *all* users of health services reduces the relative share of national income going to some of the most prestigious and politically powerful groups and institutions in society, and poses formidable political challenges to anyone who would undertake it. But such controls, which have been of demonstrated success for hospitals and, to a lesser extent, nursing homes, reduce program cost while, under some circumstances, *enhancing* access for the poor and elderly. At the same time, they reduce the general level of inflationary pressure in society, and thus benefit all citizens, as consumers and taxpayers.[15]

Conclusions

For those most actively involved in their formulation and adoption, over the longest period of time, Medicare and Medicaid were clearly and consciously to be incremental steps on the path towards universal, comprehensive health care benefits. Progress has been forestalled by the cost increases that, at least in some part, were engendered by those programs themselves, by the failure to contain program costs, and by the increasing scarcity of resources for social investment. Indeed, the path may have begun to wind backwards.

Thus, it appears we can no longer afford increasing equity in access to health care services, perhaps not even the degree of equity we have attained. At the practical, budgetary level, the annual cost increases in Medicare and Medicaid exceed the combined total expenditures of all other health service programs supported by the Department of Health and Human Services. Discretionary funds that might go to Urban or Rural Health Initiatives, expanded mental health services, or increasing the supply of manpower to under-served areas are consumed by price inflation in existing entitlement programs.

We can in fact afford increased equity, if we are willing to pay the political costs. It is clear that Medicare and Medicaid costs can not be effectively controlled by themselves without a still unacceptable degree of benefit reduction. Only true cost containment can make service expansion possible; at the same time, only true cost containment will work in the long run. From a technical perspective, it is easier, not harder, to control total hospital revenues or total long-term care expenditures than the expenses of a single payor. Politically, of course, it is much more difficult. But the choice is clearly a political one.

References

[1] Boulding E. *The concept of need for health services.* MILBANK MEMORIAL FUND QUARTERLY 1966 October.

[2] Davis K, Schoen C. HEALTH AND THE WAR ON POVERTY: A TEN-YEAR APPRAISAL. Washington: The Brookings Institution, 1978;42.

[3] Aday L. *The impact of health policy on access to medical care.* MILBANK MEMORIAL FUND QUARTERLY. 1976;215.

[4] Kasper JA, Walden DC, Wilensky GR. *Who Are the Uninsured? Data Preview 1,* NATIONAL HEALTH CARE EXPENDITURES STUDY, National Center for Health Services Research, U.S. Department of Health and Human Services.

[5] U.S. Congress, Congressional Budget Office. EXPENDITURES FOR HEALTH CARE: FEDERAL PROGRAMS AND THEIR EFFECTS. Washington: Government Printing Office, 1977.

[6] Feder JM. MEDICARE: THE POLITICS OF FEDERAL HOSPITAL INSURANCE. Lexington, MA: Lexington Books, 1977;53–101.

[7] Holahan J, Scanlon W, Spitz B. *Restructuring Federal Medicaid Controls and Incentives.* An Urban Institute Paper on Health. Washington: The Urban Institute, 1977.

[8] Fisher CR. *Differences by age groups in health care spending.* HEALTH CARE FINANCING REVIEW. Spring 1980;65.

[9] Hearn R, Altman D. *Expand Public Hospitals and State and Local Health Departments.* Prepared for 1981 Commonwealth Fund Forum (in press).

[10] U.S. Department of Health and Human Services, Public Health Services, Office of Health Resources Opportunity. *Health of the Disadvantaged, Chartbook-II.* Hyattsville, MD: Health Resources Administration, 1980;68–69.

[11] Rosenblum RW. *The Evolution of Health Care Financing Programs for the Aged in the United States: The Economic Impact on the Aged Consumer.* Unpublished doctoral dissertation, Columbia University, 1979.

[12] Vladeck BC. UNLOVING CARE: THE NURSING HOME TRAGEDY. New York: Basic Books, 1980;77, 95.

[13] Scanlon W. *Aspects of the Nursing Home Market: Private Demand, Total Utilization, and Investment.* Urban Institute Working Paper. Washington: The Urban Institute, 1978.

[14] Willemain TR. *Nursing home levels of care: Reimbursement of resident-specific costs.* HEALTH CARE FINANCING REVIEW. 1980;47.

[15] Finley JE, Vladeck BC. *Statement on Medicaid Cost Containment,* to the Subcommittee on Health and The Environment, Committee on Energy and Commerce, U.S. House of Representatives, June 5, 1981.

Equity of Access to Medical Care: A Conceptual and Empirical Overview

Lu-Ann Aday, Ph.D.*
Ronald M. Andersen, Ph.D.*

Introduction

The ethical considerations in the conceptualization and measurement of equity of access to medical care, the curent empirical profile of equity in the United States, and the implications of these conceptual and empirical issues for future health policy dealing with the "equity of access" objective are the focus points of this report.

Implicit in this objective of equity are three assumptions that serve as a starting point for examining the ethics of equity: (1) Health care is a right; (2) The resources for allocating health care are finite; (3) Health policy should be concerned with the design of "just" mechanisms for allocating scarce health-care resources. These assumptions form the basis for much of the conceptual and methodological interest in equity. Selected literature dealing with these assumptions is summarized to provide an overview of the normative context from which fundamental questions regarding the "ethics" of the equity concept have emerged in recent years.

*Center for Health Administration Studies, Graduate School of Business, The University of Chicago, Chicago, Illinois.

This study, which was originally prepared for the Commission, is reprinted with permission from MEDICAL CARE, Vol. 19, No. 12 (Supplement), 1981.

March 1981

The original research on which some of the data reported here were based was supported by grants from The Robert Wood Johnson Foundation (Princeton, New Jersey) and Contract # HRA 230-76-0096 from the National Center for Health Services Research (Hyattsville, Maryland).

Right to Health Care. In the medical ethics literature, a great deal of attention has been devoted to examining the concept of the "right to health." Blackstone,[1] for example, explores the constitutional and moral bases for this right in terms of equal protection, minimum welfare, and other "justice" criteria. Sidel[2] and White,[3] however, point out the importance of distinguishing the "right to *health*" from the "right to *health care*." They argue that a variety of factors may contribute to health that are beyond the power of the state to ensure it as a basic "right," e.g., personal behavior that leads to ill health. The "right to health care," however, implies *access* to medical care services, which may or may not lead to "health" per se. Carlson[4] and Illich[5] argue, for example, that the modern medical care system may in fact be ineffectual or even deleterious in effecting "health" in patients who use it.

Blumstein and Zubkoff,[6] Mechanic,[7] and Sparer,[8] however, point out that in general "the right to health care" is concerned with ensuring that everyone has a right of access to services and that this right is an important normative goal of the health-care system. McGarvey[9] further distinguishes the "right *to* health care" from the "right *in* health care" in that where the former is concerned with "access," the latter deals with the character of the individual interaction between the patient and the physician, once entry (access) is gained. Both are or should be proper concerns of medical ethics, but it is the right *to* medical care that is one of the most fundamental *normative* starting points for the consideration of access per se.

Finitude of Health Care Resources. Gustafson[10] argues that when the finitude of resources to satisfy human needs is realized, one's own self-interest is secured or greater attention is paid to applying principles of distributive justice for the common good. The choice between these alternatives is, in a sense, a moral one. The shortage of a desired good relative to the demand for it is expressed in the economic marketplace through high prices for that scarce resource. The finitude of resources for securing the right to health care is correspondingly visible in the rising cost of medical care, borne by individuals and society, as the demand for and price of high-technology care continue to increase.[11-13]

The recognition of the finitude of health-care resources and the means for allocating them serves to establish "real world" parameters on how or whether the "moral" right or goal of access to medical care can actually be realized. Further, to the extent resources are devoted to the realization of the access goal, other health-care system goals such as cost containment and quality may well be compromised. These tradeoffs are also of essential concern in understanding the ethics of health policy-making overall.

"Just" Allocation of Finite Health Care Resources. Much of the debate concerning means for assuring the "right to health care" concerns whether health care differs from other commodities purchased and consumed in the marketplace, the applicability of the "market" model to distributing health care, and the "fairness" of the various mechanisms for allocating finite health-care resources.[7,12,14–16]

At the heart of considering the concept of equity of access is the question: What criteria define a "just" system of allocating this scarce resource? The relevance of variant criteria to current conceptual and methodological approaches to evaluating equity is the focus of the following section.

Conceptual and Methodological Issues in Equity of Access

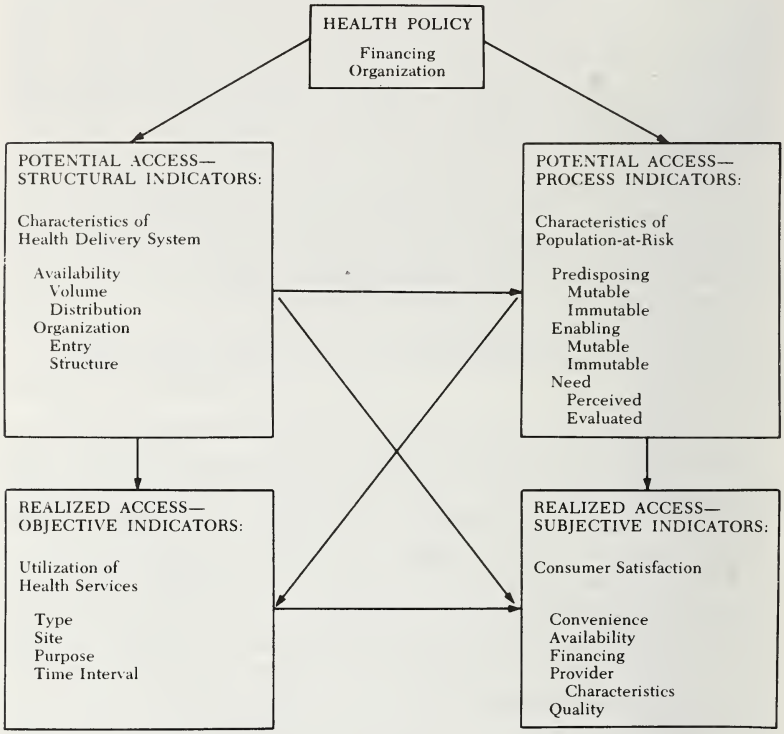
A variety of approaches to defining and measuring access to medical care have appeared in the literature in recent years. Some researchers have emphasized the overall availability (supply) of services; others have focused on the characteristics of the people who may potentially avail themselves of these services, considering factors such as their income levels or insurance coverage. More recently the focus of access studies has shifted to intermediate outcome measures, such as the rates at which services are actually used or how satisfactory consumers perceive their care to be.

The authors have developed a framework for studying access to medical care that attempts to integrate these diverse approaches.[17,18] In this framework, access is defined as "those dimensions which describe the potential and actual entry of a given population group to the health-care delivery system."[19] Figure 1 shows an outline of the elements of this framework.

Further, in this model *equity* of access to care is said to exist "when services are distributed on the basis of people's need for them . . . Inequity is suggested, however, if services are distributed on the basis of demographic variables, such as race, family income, or place of residence, rather than need."[19] In terms of Figure 1, variation in the use of services as a function of the need component or associated age and sex correlates within the predisposing variables represent equity. Variation that is a function of the availability of services or how they are organized or other predisposing or enabling characteristics of the individuals themselves means that services are not equitably distributed.

In the discussion that follows, alternative approaches to measuring equity are introduced with particular reference to how they might be put into operation within the access framework outlined in Figure 1 and the applicability of these respec-

FIGURE 1. Framework for the Study of Access



tive criteria for effecting a "just" allocation of health-care resources.

Concepts of Equity. Outka[20] explores the relevance of a number of criteria of social justice to the "equal access" objective. They are: (1) merit or desert, (2) societal contributions, (3) contribution to satisfying whatever is freely desired by others in the open marketplace, (4) needs, and (5) similar treatment for similar cases. These concepts serve as the starting point when considering variant criteria for measuring equity and the ways in which other conceptions or interpretations of justice modify or extend these.

Merit. Outka argues that merit, or who is *deserving* of care, is not an appropriate criterion to apply when considering "just" means for allocating health-care resources because (1) the idea of justice is not fully defined by the merit criterion alone, and (2) health crises generally cannot be based on merit because they occur beyond an individual's power to control or predict them.

Outka does describe examples in which it can be argued that individuals are to some degree responsible for their medical problems; e.g., an overweight person who had a heart attack,

a man with lung cancer who smoked cigarettes for over 40 years, or a football player who suffered a concussion.

A meaningful application of this merit criterion in terms of the overall access objective, however, may be to examine those predisposing factors, such as attitudes or health behaviors, that place individuals at greater risk of needing health care. Examples of these are people with the smoking or exercise habits in the previous example, as well as persons in age and sex categories for which the biologic givens of aging or anatomy predispose them to "undeserved" health risks. These criteria are, in fact, applied in health education, screening, risk evaluation programs, and in insurance policies with reduced premiums for people with good health habits (those who do not smoke or drink, for example).

Societal contribution. The societal contribution criterion is based on a utilitarian concept of justice that considers the greatest good for the greatest number. Outka points out that those who contribute least to the general welfare (such as the elderly or disabled) may be ones for whom access to health-care support services are the most critical.

In our society, occupational and educational statuses are the criteria frequently used (either implicitly or explicitly) to assess whether an individual's contribution to the common good is worthwhile. The differential distribution of employment-related health-care benefits, for example, may be viewed as a manifestation of the resources that society is willing to devote to sustaining the health of differentially valued societal "contributors." Professional and occupational statuses cannot entirely determine whose "right to health" is most fully guaranteed, however. Medicaid and Medicare, for example, were enacted to provide coverage for individuals whose contributions to society may well be the least—the poor and the elderly.

Supply and demand. The supply and demand principle of resource allocation emphasizes the consumer's free choice of preference among alternatives and the distribution of services on the basis of the unrestricted (laissez-faire) operation of these forces. Outka and others, however, argue that the application of these principles to the distribution of health-care resources is inappropriate primarily because the consumer has much less choice when deciding whether to purchase this type of service. Outka, for example, points out that "when lumps appear on someone's neck, it usually makes little sense to talk of choosing whether to buy a doctor's service rather than a color television set."[20] Mechanic[7,12] also argues that physicians, rather than consumers, make the major decisions about which services should be purchased. The consumer has little information or knowledge of the appropriate "commodity" to buy when he or she is seriously ill. Navarro,[15,21,22] in fact, has made the case that the health-insurance industry and other powerful corporate inter-

ests, rather than either physicians or consumers, make the major decisions about the kinds of health care to be delivered and consumed in the health-care marketplace.

Economists are most apt to be proponents of the unrestricted operation of the free market forces of supply and demand for the allocation of health-care resources.[23,24] Programs should be designed, they argue, to encourage rather than restrict the consumer and the physician when deciding on the amount and price of health care.[14,25] This policy perspective tends to encourage the development of competitive models of health care delivery, such as HMOs, or the provision of income maintenance rather than health-care specific subsidies. The laissez-faire principles in this approach conform closely to Nozick's entitlement concept of justice, in which free choice among alternatives and entitlement based upon those historic choices are the main criteria on which a just allocation of resources is based.[26]

The access framework outlined earlier summarizes the supply (delivery system characteristics) and demand (population characteristics) factors that determine the price and quantity of services consumed in the health-care marketplace. A key health policy concern in examining these forces or influences is the extent to which there should be intervention to directly regulate the behavior or operation of these determinants of who obtains how much care. Most contemporary medical economists argue that some "basic minimum" care should be assured by society; however, market forces should be freer to operate in allocating services beyond this minimum.[6,25] Fein[26] argues for introducing new incentives into the existing market (e.g., centralized ceiling on health expenditures, prepaid forms of organizing care) and allowing the system to respond to these incentives.

Need. Outka[20] contends that "essential needs," in fact, constitute a much more appropriate basis for the allocation of health-care resources than do merit, societal contribution, or supply and demand criteria alone. The previous discussion has suggested that because of the ethical uniqueness of allocation decisions concerning medical care, these latter criteria have, in fact, either been tempered, transcended, or ignored in many major health policy-making efforts.

As mentioned earlier, this criterion based on need has been the primary one used by the authors to describe an equitable system of health-care resource distribution. Rawls's concept of social justice, based upon the allocation of resources to those individuals in the society who have the greatest need, lends support to this approach, as do the arguments of those who contend that medical care is a unique commodity for which the typical market force of merit-based allocative criteria may not be appropriate.[27,28] For example, Fein has argued that we should

"reject market mechanisms that might allocate resources to and within the health sector, in part, because market results are at variance with our values."[26]

The need criterion may be made operational in terms of a minimum service (e.g., annual Pap smears for women in childbearing years), the identification of pockets of observed morbidity (e.g., lead paint poisoning), or measures of the fit between observed need and obtained services (e.g., use of services relative to experienced symptoms or disability). It is this latter utilization criterion based on need that has received the greatest attention in our own approach to measuring equity.

Implicit in the development of these need criteria are some assumptions regarding the appropriate level of care relative to experienced need. The development of these criteria often requires consumers or providers of care to make value judgments on how much of what service is required.[6] In addition, methodological criticisms have been made of the techniques used to put this approach into operation.[19] Nonetheless, it is a perspective that has received considerable attention as a basis for allocating health-care resources and will be of particular interest in some of the empirical analyses of access that follow.

Similar treatment for similar cases. This criterion can be explored in the context of the access framework described earlier when differences in the rates at which people report having a regular source of medical care, the convenience of this source, or the satisfaction with this care are examined. Income, race, or whether a person is covered by Medicaid or private insurance should not be used to determine the medical care received by people who must enter different doors or be "treated" differently by the health-care system. This criterion, in particular, focuses on sociodemographic subgroup comparisons. Fein refers to distributional equity on this basis as "vertical equity."[26]

Measurement of Equity. The preceding discussion has described the primary ways in which the equity of access concept has been defined when formulating health policy decisions. Table 1 summarizes how each of these approaches can be put into operation, with particular reference to the access framework outlined earlier.

In our application of the access framework to measure the equity concept, we have tended to focus on the last two concepts of equity that are based on need and similar treatment for similar cases. In particular, we have emphasized the development of need-based measures of utilization and comparisons of race, income, residence, and other groups for whom similar treatment for comparable levels of need are expected.[19]

As Table 1 suggests, however, the access framework can actually be applied when putting into operation each of the approaches used to measure equity just described. Personnel

TABLE 1. Application of Access Framework to the Measurement of Equity Concepts

Dimensions of Access	Empirical Indicators of Access	Concepts of Equity				
		Merit	Societal Contribution	Supply and Demand	Need	Similar Treatment for Similar Cases
Characteristics of health delivery system						
Availability	Personnel/facility-to-population ratios			X		
Organization	Structure and convenience of medical care facilities					X
Characteristics of population-at-risk						
Predisposing	Age		X		X	
	Sex				X	
	Race					X
	Education		X			
	Occupation		X			
	Health care attitudes	X				
	Health care behaviors*	X				
Enabling	Rural-urban place of residence			X		X
	Family income		X	X		X
	Structure and convenience of regular source of care					X
	Type and extent of insurance coverage		X	X		
Need	Disability days	X			X	X
	Symptoms	X			X	X
Utilization of health services						
Physician				X	X	X
Hospital				X	X	X
Dentist				X	X	X
Consumer satisfaction						
General satisfaction						X
Visit-specific satisfaction						X

* This variable did not appear in the original access framework.

and facility-to-population ratios, for example, are useful in exploring the operation of supply and demand factors. The convenience and characteristics of places people go for medical care provide data on whether there is differential "treatment" of individuals, depending upon where they happen to go for services.

Age is a correlate of need as well as a basis for defining actual or potential societal contributions; e.g., children, working-age adults, and the elderly. Sex, similarly, is often considered, rather than need; race serves to identify groups for whom access differences may well indicate differential treatment. As mentioned earlier, education and occupation are used as proxy indicators of the contributions individuals make to society. Health-care attitudes and behaviors reflect predispositions toward health and health care that may serve as a basis for deciding whether certain individuals merit the allocation of scarce health-care resources. Health habits, such as smoking, diet, and exercise, did not appear in our initial access framework. They

are, however, ultimately important determinants of the need for and utilization of health-care resources and may be appropriately considered among the factors predisposing individuals to eventually seek services.

The rural-urban distribution of services is important from the point of view of both the supply and demand and similar treatment concepts of equity. According to the similar treatment norm, for example, where a person lives should not determine the adequacy of medical services available. Family income serves as an indicator of the "value" individuals have in society, as well as a basis for determining the operation of supply and demand factors. Differential access by family income, for example, may suggest that differential purchasing power rather than need per se determines who receives care.

Insurance coverage is a mechanism for affecting the operation of supply and demand factors and ultimately the price of services to those whose societal contribution is great (e.g., employees) or whose previous (elderly) or potential (children) societal contributions merit attention. The characteristics and convenience of places people usually go for medical care indicate whether individuals seem to have different experiences with different types of providers (e.g., longer waiting times in hospital emergency rooms, shorter visits with general practitioners compared with specialists) and what these differences might mean in terms of the "equitable" allocation of health-care resources.

Disability days and symptoms are empirical indicators of underlying medical morbidity. They may be used as statistical controls to determine whether access differentials persist. Morbidity is also important in determining whether certain individuals merit the allocation of health-care resources (e.g., life-threatening or potentially disabling illnesses may signal those conditions most deserving of attention).

Utilization measures are essential in supply and demand formulations of equity. Utilization indexes based on need are also important when need is used as the basis for resource allocation. That is, we expect similar utilization rates for individuals whose needs are similar.

Satisfaction levels indicate whether people have experienced differential treatment. In general, the more inconvenient or costly people's experiences with the health-care system, the less satisfied they are.

The operational indicators just reviewed are the ones that are oftentimes available from empirical studies of access.[29] They are more or less relevant, as indicated by the previous discussion, to measuring a particular concept of equity. Further, one indicator may be applied in more than one approach to the equity concept.

Methodological Problems in Measuring Equity. A number of methodologic problems that plague any effort to empirically measure equity should be noted, however. They are, in fact, limitations of health services research methodology in general. Their particular relevance to the measurement of the equity of access objective is summarized here.

Empirical research has failed, for example, to demonstrate a necessary relationship between gaining access to medical care and health. Many factors (e.g., genetics, environmental pollution, health habits, occupational stresses) other than health care per se can affect one's overall health status. There is some evidence, as cited earlier, that health "care" may, in fact, lead to poorer rather than better health. The equity of access objective, however, is concerned with individuals gaining the right to "take the chance" that obtaining such care will benefit rather than harm them.

The "science" of medicine has not been able to establish universally accepted criteria for what is the most appropriate care for any particular illness. Standards of the appropriateness of access for a given level of need are still in large measure judgmental. Some critics have argued, in fact, that consumers may be as qualified as medical professionals to judge whether it is appropriate for them to seek scientific medical care for particular illnesses. Consumer satisfaction levels and attitudes are equally relevant criteria for evaluating the "success" of the medical profession in achieving their curing and caring objectives.[30]

As Mechanic[31] and others have pointed out, much of the variance in understanding the utilization of or access to medical care services remains unexplained by large-scale empirical studies in this area. One may argue that empirical work to date has done a poor job of testing and measuring the equity of access concepts or that current conceptions do not adequately capture the full range of factors that may influence who gets care and why. The extent to which health-services research can help determine health policy and the ethical choices associated with it is bounded by the success with which it can empirically model these choices and their implications for the individuals and the health-care system.[32] To the extent a great deal of the variance in access remains unexplained, the precise profile of equity may correspondingly be inadequately described.

Other methodologic issues used to empirically measure equity concern the direction and magniture of observed relationships. For example, does not having a regular source of medical care mean people seek services less often, or do people who are healthier have less need to establish a relationship with a regular provider? Or how much difference by income or race is "enough" to establish that there is an access "problem"? Also, might the ethical importance of certain determinants vary

depending upon the particular type of service being considered, e.g., emergency medical care versus cosmetic surgery? Further, Sloan[33] points out that different generalizations about equity may be the result of the particular concept or empirical definition of equity that is used.

Methodologic problems, therefore, do limit the generalizations that can be made about the profile of equity that exists at any point in time. In the section that follows the most current empirical information on the status of the equity goal is summarized, taking into account some of the limitations just described.

Trends in the Profile of Equity

Since the late 1950s, national data on the profile of access to medical care have shown increasing improvement for traditionally disadvantaged groups.[19,34,35] The advent of large-scale public financing programs (Medicare and Medicaid) and programs targeted to low-income populations (community health centers, children and youth programs, maternal and infant care) have been credited with these improvements.

Since the mid-1970s, however, there have been substantial changes in the overall economic and political climate of the country. Accompanied by programmatic changes, they may well temper the progress toward equity observed in previous decades.

We have witnessed in recent years, however, an era of double-digit inflation in tandem with high rates of unemployment. Taxpayers have begun to more vigorously protest the amounts they must pay into the public coffers. Local, state, and Federal lawmakers see cost containment and balanced budgets as increasingly important, but difficult, goals to achieve.

There is an increasing skepticism about what "big government" can accomplish and more blatant criticism of the perceived self-interest and inefficacy or even incompetence of the medical as well as other (e.g., legal) professions.

The poor economic health of the nation and general conservative attitude toward the efficacy and efficiency of existing programs have certainly had an impact on the health sector, as manifested by cutbacks in the Medicaid programs in a number of states, the closings of inner-city public hospitals in many large cities, the inability of Congress to reach agreement on a national health insurance proposal, and the increased burden on individual families who lose health-insurance benefits during periods of unemployment.

Under these conditions, optimization of the goals of equity of access, cost containment, and quality is difficult to attain. There is some evidence that the public financing program credited with substantial improvements in access for low-income

persons and nonwhites may have, in fact, significantly contributed to the escalation in the cost of medical care in recent years. Further, though utilization rates may be higher now for these groups than they ever were in the past, where they go for care (e.g., hospital emergency rooms, Medicaid mills, physicians who are not specialty-board certified) suggests that the quality of the care available to them may be less than optimum or, in fact, poor in many cases.

In the following section the most recent national data on the equity of access profile are presented in an effort to determine where we now stand. In particular, we examine whether the generally positive trends of the previous decades are continuing, whether they are apt to be maintained, and in which areas existent gains are most apt to be compromised as a result of the changes in the overall economic, political, and programmatic climate of the nation just described.

Current Survey Data on Equity. The primary sources of the data presented here are the national surveys of medical care utilization conducted by the Center for Health Administration Studies and the National Opinion Research Center (CHAS-NORC) at The University of Chicago,[19,34] data from the National Center for Health Statistics Health Interview Surveys (NCHS-HIS), and preliminary data from community survey evaluations of The Robert Wood Johnson Foundation Community Hospital Program (RWJF-CHP) carried out by the Center for Health Administration Studies in 1978 and 1979.[36,37]

The concepts of equity that were discussed earlier and are particularly emphasized in the data presented here are the need and similar-treatment-for-similar-case criteria. Poor–nonpoor and age-group comparisons are especially emphasized, as are adjustments for variants in need for services. Data are presented on both indicators of potential (e.g., insurance coverage) and realized (e.g., utilization) access to services.

Table 2 summarizes trends for five access indicators: percent having third-party (public or private) coverage, percent of women with live births who saw a doctor during the first trimester of pregnancy, hospital admissions per 100 persons per year, doctor visits per person per year, and the percent seeing a dentist in the year. Data are presented for five different points in time over an approximately 25-year period—1953, 1958, 1963, 1970, and 1976—covered by the CHAS-NORC national surveys.

The table shows the ratio of each indicator value for the poor compared with the nonpoor. A score greater than 1 means that the poor are in a better position (have higher scores) than the nonpoor. A value of less than 1 means they are in a less favorable position. A value of 1.00 means 100 percent parity (or agreement on the actual scores) between the two groups.

TABLE 2. Selected Indicators of Access to Medical Care of the Poor* Compared With the Nonpoor: From the CHAS National Surveys† (1953, 1958, 1963, 1970, 1976)

Access Indicators	Ratio Score: Poor/Nonpoor‡				
	1953	1958	1963	1970	1976
Per cent with third-party coverage of medical expenses	0.70	0.65	0.73	0.87	0.84
Per cent of women with live births who see physician in the first trimester of pregnancy	0.56	0.83	0.67	0.74	0.73
Hospital admissions per 100 persons per year	1.02	1.08	1.05	1.36	1.29
Physician visits per person per year (excludes phone calls)	0.70	—	0.98	1.32	1.12
Per cent seeing dentist within a year	0.37	—	0.48	0.57	0.60

* Source: References 19, 34, 38, 39, 40, 41.

† Definitions of poor based on total family income of under: (1953) $3,000; (1958) $3,500; (1963) $4,000; (1970) $6,000; (1976) $8,000.

‡ Score less than 1.00 indicates poor have less access; score greater than 1.00 indicates poor have greater access.

Most broadly conceived, a value of 1 on the scale score axis indicates a situation of equity by the similar-treatment criterion. Some variance from this norm may well be due to real differences in need, which is a justifiable basis (according to the need criterion) for discrepancies among groups for certain of the indicators examined here (e.g., hospital admissions, physician visits).

According to Table 2 the poor continue to be at a relative disadvantage compared with the nonpoor with respect to the proportion having some form of third-party insurance coverage. Though with the advent of Medicare and Medicaid their situation has improved considerably since the mid-1960s, the proportion of the poor having any form of coverage (public or private) is still less than the nonpoor. Further, in recent years the poor do not appear to be gaining on the rest of the population.

The proportion of low-income women seeing a doctor during the first trimester of pregnancy continues to be lower than that of higher-income mothers. There is evidence of substantial gains between 1953 and 1958, which have not been fully maintained in the past two decades. There was little change between 1970 and 1976 in the relative position of poor mothers. They continued at 73–74% of parity compared with high-income mothers.

The poor have generally had higher rates of hospital admissions than the nonpoor over the past 25 years. One reason is that the poor tend to have much more serious, especially chronic, illnesses compared with the nonpoor. A major increase in the relative number of admissions of the poor compared with the nonpoor occurred in the post-Medicare-Medicaid period (from 5% more in 1963 to 36% more in 1970). This relative increase leveled off over the next 5-6 years, however.

Especially dramatic gains of the poor relative to the nonpoor may be observed in terms of the mean number of physician visits reported over the 25-year period. In the early 1950s the poor averaged 70% as many visits as the nonpoor. From that time on their relative position improved; of major impetus were the federal financing initiatives during the mid-1960s. By 1970 the poor were averaging considerably more visits than the nonpoor (approximately one-third). By 1976 this relative advantage appeared to decline somewhat.

The relative position of the poor in terms of the proportion who have seen a dentist during the year improved substantially during the 1950s and 1960s. However, less improvement is noted during the 1970s and their rate of use remains considerably below that of the nonpoor (about 60% of parity).

In summary, Table 2 reinforces our conclusions from earlier analyses of the CHAS surveys that access of the poor rela-

tive to the nonpoor has improved significantly since the early 1950s. Substantial inequities persist, however, in the areas of dental services, prenatal care, and adequate third-party coverage; there have been no signs of these gaps narrowing in recent years. Further, dramatic increases observed in the 1950s and 1960s with respect to rate of hospital and physician use by the poor compared with the nonpoor did not appear to continue into the 1970s.

Table 3 (from NCHS) provides even more recent data, including 1978 and 1979, to illuminate whether these trends are continuing. Also, the use of a second major data source is important for confirmation because of the complexity of the access concept and the rudimentary state-of-the-art concerning its measurement.

The data in Table 3 tend to confirm the broad profile of equity on the part of the poor and nonpoor observed in Table 2. The poor continue to have a considerable disadvantage compared with the nonpoor with respect to dental service use. The leveling off of gains on this indicator between 1963–64 and 1976 appears to be sustained through 1978.

The NCHS hospital discharge data show a pattern similar to the CHAS-NORC hospital admission rate estimates—dramatic gains for the poor in the mid to late 1960s with a leveling off in the 1970s.

The NCHS data—as do CHAS-NORC* data not reported here—suggest that the proportion of the poor and nonpoor contacting a doctor at least once is approaching parity. The NCHS mean physician-visit data similarly indicate there has been a sustained increase in the relative position of the poor through the 1960s; in the 1970s the rate of this increase began to decline to some extent.

To interpret more adequately the equity of the differences observed between the low- and high-income groups, controls for the variant levels of illness experienced by the respective groups and their differential age composition should be attempted. For example, the poor have traditionally reported more days of disability and higher rates of chronic illness than the nonpoor.[42] We have observed from recent overall mean physician visit and hospital use data that the poor now, in fact, use more services on the average than do the nonpoor. This appears to be a move in the right direction to achieve equity. Tables 4–6 present several alternatives to estimating equity between the high- and low-income groups, computing physician use rates adjusted for need and age. These indicators represent an effort

* The CHAS-NORC and NCHS-HIS estimates of physician visits differ slightly in that the former excludes phone calls, while the latter includes them.

TABLE 3. Selected Indicators of Access to Medical Care of the Poor* Compared With the Nonpoor: From the NCHS National Surveys† (1964, 1973, 1976, 1978, 1979)

Access Indicators	Ratio Score: Poor/Nonpoor‡				
	1964	1973	1976	1978	1979
Hospital discharges per 100 persons per year	1.10	1.52	1.47	1.51	—
Per cent visiting a physician in past two years	0.88	0.96	0.97	0.99	—
Physician visits per person per year	0.94	1.14	1.17	1.21	1.22
Per cent visiting a dentist in past two years	0.57	0.67	0.70	0.70	—
Dentist visits per person per year	0.44	0.61	0.61	0.61	—

* Source: Unpublished data from the Health Interview Survey, National Center for Health Statistics.

† Definitions of poor based on total family income of under: (1964) $3,000; (1973) $6,000; (1976) $7,000; (1978) $7,000; (1979) $7,000.

‡ Score less than 1.00 indicates the poor have less access; score greater than 1.00 indicates the poor have greater access.

to refine empirically and apply the concept of equity based on need that was discussed earlier.

Table 4 summarizes the rates of use relative to need for the poor and nonpoor from the CHAS surveys in contrast to physician-determined norms of "appropriate" access. The symptoms-response ratio reflects whether individuals in different age groups see a doctor more or less often for symptoms experienced than a panel of physicians thought was warranted.[19] A zero score on the index means individuals saw a doctor as often as the doctors thought necessary. A minus score means doctors were seen less often and a plus score indicates doctors were seen more often than doctors thought appropriate.

TABLE 4. Symptoms-Response Ratio for the Poor* Compared With the Nonpoor: From the CHAS National Surveys† (1963, 1970, 1976)

	Symptoms-Response Ratio Score‡		
Family Income Level	1963	1970	1976
Poor	−28.00	−14.00	−5.00
Nonpoor	−12.00	−11.00	−6.00

* Source: Reference 19.
† Definitions of poor based on total family income of under: (1963) $4,000; (1970) $6,000; (1976) $8,000. In each year approximately one-fourth of the population was defined as poor.
‡ A minus score indicates group saw a doctor less often for symptoms experienced than panel of physicians judged appropriate. A plus score would indicate group saw a doctor more often than panel judged appropriate.

Table 4 provides a very encouraging view of the access gains of the poor between 1963 and 1976. The tendency for low-income individuals to see a doctor much less often (28% less often in 1963, for example) than the physicians thought necessary has changed dramatically over time. According to the symptoms-response ratio for 1976, the poor and nonpoor had very similar profiles of appropriate contact for symptoms—5% and 6% less often than the physicians judged appropriate, respectively. However, there is still a tendency for both low- and high-income individuals to see a physician less often for symptoms than clinically-based norms suggest is appropriate.

Tables 5 and 6 provide the NCHS alternative need-adjusted physician visit measures for years 1964 through 1979. Table 5 is a modified use-disability index that summarizes the ratio of mean physician-visits to mean disability days for the respective

age and income groups.† A higher score on the ratio means that the group had more visits relative to days of disability than those with a lower score. This ratio may be considered a physician-use indicator adjusted primarily for acute illness because the majority of the disability days reported are associated with acute conditions. In contrast, the basis for the data reported in Table 6 is mean physician-visit estimates for the respective groups adjusted for the proportion with chronic illness in each of the groups.‡

According to Table 5, the poor have not achieved parity with the nonpoor in terms of a very general index of visits relative to reported days in bed or of limited activity. Poor children and young adults appear to have come closest to parity. This index shows considerable fluctuation rather than consistent improvement for the poor from the early 1960s through the late 1970s. Indeed, the ratios from the latest year (1979) are very similar to the initial ones shown for 1964.

According to the index adjusted for chronic illness in Table 6, there appears to have been substantial improvement in the status of the poor of all age groups since the early 1960s. The relative improvement for poor children—from 57% of parity (0.57) in 1964 to virtual parity (96% or 0.96) in 1979—is particularly noteworthy. As all of the poor groups approached or exceeded parity in the mid-1970s we observe fluctuations in the trends, but these fluctuations should be interpreted with caution. There is some suggestion that the gains by the poor in certain age groups cannot be maintained.

In sum, the three indices of physician visits adjusted in various ways for illness (Tables 4–6) provide a somewhat mixed picture. The symptoms-response ratio (Table 4) and visits adjusted for chronic illness (Table 6) shows the poor making consistent progress from the early 1960s through the mid-1970s. Information from the later 1970s on the chronic illness measure fails to substantiate systematic and sustained gains for the poor in all age groups. It might be argued, according to the need concept of equity, that the higher scores for the poor are not justified. What would be disturbing, however, is the possibility that we are documenting the beginnings of a reversal of the earlier favorable trends; this could result in the poor systematically losing ground and regressing to their traditional disadvantaged position. The use-disability ratio (Table 5), like the chronic illness adjustment, fails to show systematic gains for the poor in the late 1970s but also does not support the

† This modified ratio differs from the original in that it is not limited to individuals who had at least one day of disability in the year. See Reference 19 for a fuller description of how the original index was constructed.

‡ These adjustments are explained in HEALTH: UNITED STATES, 1980.[35]

view that the poor made systematic gains in the 1960s and early 1970s.

Tables 7–10 provide comparisons among 12 different communities and the national average for selected access indicators. These data are based on access surveys conducted in representative communities in which there was a community hospital-sponsored group practice funded by The Robert Wood Johnson Foundation as a part of its Community Hospital Program (CHP).[36,37]

These data were for the most part collected in late 1978 and early 1979.† The purpose of presenting these findings is to try to discover if there are continuing pockets of inequity among certain groups or types of communities that may not be apparent from the aggregate national statistics just reviewed. The communities in the Community Hospital Program represent a range of community types—high- and low-income, metropolitan (densely populated) and nonmetropolitan (sparsely populated), regionally diverse, etc.

Tables 7–9 show the ratios of the scores for the community compared with the national averages on the access measure of interest. A value of 1 means conformity with the national average, a value of less than 1 means the community score is lower, and a value of more than 1, the community score is higher than the national average. The scores for various age groups in the respective communities are displayed in each table.

Assume that a score of ±10% of the national average (range of 1.1–0.9) indicates that the community and the nation as a whole conform relatively closely. Most age groups within each of the CHP communities are near the national average in terms of the proportion of residents who report having a regular source of care (Table 7). Two communities from the Western region (C and H) are exceptions, with all age groups except the elderly showing limited access to a regular source of care. Young adults, 18–34 especially, were less likely to report having a regular place to go for medical care than was generally true nationally. In contrast, the elderly in these communities were within the ±10% band. We may be observing here the equalizing impact of Medicare at the community level.

The proportion of people with some form of public or private third-party coverage was also lower in Community C—a low-income inner-city area—than was the case nationally (Table 8). Among the other communities there was greater correspondence in the rates of coverage to the national average. However, it appears that the high-income communities tend to have a higher proportion of their residents reporting coverage than the lower-income communities.

† These data are preliminary and subject to change.

TABLE 5. Physician Visits/100 Disability Days for the Poor* Compared With the Nonpoor, By Age Group: From the NCHS National Surveys† (1964, 1973, 1976, 1978, 1979)

Age Group	Ratio Score:Poor/Nonpoor‡				
	1964	1973	1976	1978	1979
Under 17 years	0.78	0.72	0.79	0.77	0.78
17–44	0.69	0.70	0.72	0.74	0.74
45–64	0.49	0.47	0.55	0.60	0.46
65 and over	0.62	0.74	0.72	0.58	0.61

* Source: Unpublished data from the Health Interview Survey, National Center for Health Statistics.

† Definitions of poor based on total family income of under: (1964) $3,000; (1973) $6,000; (1976) $7,000; (1978) $7,000; (1979) $7,000.

‡ Score less than 1.00 indicates the poor have fewer visits per 100 disability days than nonpoor. Score greater than 1.00 indicates the poor have more visits per 100 disability days than the nonpoor.

TABLE 6. Poor to Nonpoor Physician Visit Ratios* Adjusted for Chronic Illness by Age Group: From the NCHS National Surveys†: (1964, 1973, 1976, 1978, 1979)

Age Group	Ratio Score Poor/Nonpoor‡				
	1964	1973	1976	1978	1979
Under 17 years	0.57	0.88	0.89	0.93	0.96
17–44	0.80	1.06	1.06	1.21	1.20
45–64	0.80	0.90	0.94	1.16	0.97
65 and over	0.75	0.87	1.01	0.85	0.90

* Source: Unpublished data from the Health Interview Survey; References 43, 44.

† Definitions of poor based on total family income of under: (1964) $3,000; (1973) $6,000; (1976) $7,000; (1978) $7,000; (1979) $7,000.

‡ Score less than 1.00 indicates the poor have fewer visits adjusted for their rate of chronic illness than the nonpoor. Score greater than 1.00 indicates the poor have more visits than the nonpoor.

TABLE 7. Proportion of Persons With a Regular Source of Care in Selected Communities (1979) Compared With the U.S. Average (1976) by Age Group*

Community	Community Characteristics			Ratio Score: Community/Nation‖			
				Age			
	Income Level†	Residence‡	Region§	Under 17 years	18–34	35–64	65 and over
A	Low	Metro	NE	1.02	1.08	0.92	0.99
B	Low	Metro	S	1.05	1.07	1.01	1.02
C	Low	Metro	W	0.79	0.71	0.83	0.98
D	Low	Nonmetro	NE	0.99	0.96	0.90	1.04
E	High	Metro	NC	1.05	1.04	1.00	0.95
F	High	Metro	NC	1.00	1.02	0.99	0.93
G	High	Metro	W	0.98	1.02	1.05	0.93
H	High	Nonmetro	W	0.84	0.75	0.81	0.90
I	High	Nonmetro	S	1.04	1.08	1.01	1.03
J	High	Nonmetro	NE	1.06	0.96	0.95	0.95

* Source: Unpublished preliminary data from Community Hospital Program Access Study, CHAS, University of Chicago; Reference 19.

† Low indicates less than $10,000; high indicates $10,000 or more.

‡ Metro indicates community is within a Standard Metropolitan Statistical Area. Nonmetro indicates community is not within a Standard Metropolitan Statistical Area.

§ Refers to major census divisions: NE = Northeast, NC = North Central, S = South, W = West.

‖ Score less than 1.00 indicates proportion in community with a regular source is less than proportion in the country as a whole. Score greater than 1.00 indicates proportion in community is greater than in country as a whole.

TABLE 8. Proportion of Persons With Third-Party Coverage of Medical Expenses in Selected Communities (1979) Compared With the U.S. Average (1976) by Age Group*

Community	Community Characteristics			Ratio Score: Community/Nation‖		
				Age		
	Income Level†	Residence‡	Region§	Under 17 years	18–34	35–64
A	Low	Metro	NE	1.04	1.05	0.95
B	Low	Metro	S	1.03	0.96	0.99
C	Low	Metro	W	0.81	0.74	0.88
D	Low	Nonmetro	NE	0.93	0.85	0.98
E	High	Metro	NC	1.12	1.12	1.10
F	High	Metro	NC	1.05	1.06	1.04
G	High	Metro	W	1.07	1.04	1.01
H	High	Nonmetro	W	0.94	0.90	0.99
I	High	Nonmetro	S	1.11	1.07	1.09
J	High	Nonmetro	NE	1.09	1.09	1.05

* Source: Unpublished preliminary data from Community Hospital Program Access Study, CHAS, University of Chicago; Reference 19.

† Low indicates less than $10,000; high indicates $10,000 or more.

‡ Metro indicates community is within a Standard Metropolitan Statistical Area. Nonmetro indicates community is not within a Standard Metropolitan Statistical Area.

§ Refers to major census divisions: NE = Northeast, NC = North Central, S = South, W = West.

‖ Scores less than 1.00 indicate proportion in community with third-party coverage is less than proportion in country as a whole. Scores greater than 1.00 indicate proportion in community is greater than in country as a whole.

TABLE 9. Proportion of Persons Visiting an M.D. Within a Year in Selected Communities (1979) Compared With the U.S. Average (1976) by Age Group*

Community	Community Characteristics			Ratio Score: Community/Nation‖			
				Age			
	Income Level†	Residence‡	Region§	Under 17 years	18–34	35–64	65 and over
A	Low	Metro	NE	1.02	0.91	0.86	1.06
B	Low	Metro	S	1.03	1.02	1.04	1.09
C	Low	Metro	W	0.99	0.87	0.96	1.01
D	Low	Nonmetro	NE	1.05	1.00	0.91	0.97
E	High	Metro	NC	1.17	0.94	1.01	0.98
F	High	Metro	NC	1.06	0.99	0.93	0.94
G	High	Metro	W	1.24	1.01	1.03	0.89
H	High	Nonmetro	W	1.15	1.02	0.98	0.94
I	High	Nonmetro	S	1.18	1.04	0.99	1.02
J	High	Nonmetro	NE	1.17	1.03	1.04	1.10

* Source: Unpublished preliminary data from Community Hospital Program Access Study, CHAS, University of Chicago; Reference 19.

† Low indicates less than $10,000; high indicates $10,000 or more.

‡ Metro indicates community is within a Standard Metropolitan Statistical Area. Nonmetro indicates community is not within a Standard Metropolitan Statistical Area.

§ Refers to major census divisions: NE= Northeast; NC = North Central, S = South, W = West.

‖ Score less than 1.00 indicates proportion in community visiting a physician is less than proportion in country as a whole. Score greater than 1.00 indicates proportion in community is greater than proportion in country as a whole.

TABLE 10. Symptoms-Response Ratio for Selected Communities (1979) by Age Group*

Community	Community Characteristics			Symptoms-Response Ratio Score‖			
	Income Level†	Residence‡	Region§	Age			
				Under 17 years	18–34	35–64	65 and over
A	Low	Metro	NE	26	34	− 3	28
B	Low	Metro	S	18	−12	− 1	24
C	Low	Metro	W	4	−11	11	8
D	Low	Nonmetro	NE	1	−15	−24	−17
E	High	Metro	NC	18	13	−17	10
F	High	Metro	NC	9	− 7	− 1	−12
G	High	Metro	W	11	−17	−10	−23
H	High	Nonmetro	W	−12	−22	−14	1
I	High	Nonmetro	S	6	7	−17	6
J	High	Nonmetro	NE	14	− 8	− 9	4
National Average				− 1	−11	− 6	0

* Source: Unpublished preliminary data from Community Hospital Program Access Study, CHAS, University of Chicago; Reference 19.

† Low indicates less than $10,000; high indicates $10,000 or more.

‡ Metro indicates community is within a Standard Metropolitan Statistical Area. Nonmetro indicates community is not within a Standard Metropolitan Statistical Area.

§ Refers to major census divisions: NE = Northeast, NC = North Central, S = South, W = West.

‖ A minus score indicates community residents saw a doctor less often for symptoms experienced than a panel of physicians judged appropriate. A plus score indicates community residents saw a doctor more often than panel judged appropriate.

The proportion of community residents who had seen a doctor within the year was within 10% of the national average for the majority of age groups in all communities (Table 9). The major exception to this trend was the generally high proportion of children in the high-income communities seeing a doctor. On the "deficit side," adults (except the elderly), particularly in the low-income communities, tend to have contact rates below the national average. The most extreme case is the adult group in Community A (a low-income inner-city area in the Northeast) with a contact rate that is 86% of the national average.

Table 10 compares the symptoms-response index in each community with the norms of appropriate response provided by the physician panel. It shows that there is considerable variance by community and age group in the symptoms-response ratio scores.† Some particular subgroups are considerably below this figure, e.g., 35–64-year-olds in Community D (a low-income relatively rural area in the Northeast), and some are considerably above it, e.g., 18–34-year-olds in Community A. Despite unique community differences, some patterns reflective of national findings do emerge[19]; higher scores among children and those in metropolitan areas, and lower scores for adults and individuals in nonmetropolitan areas.

In summary, Tables 2–10 suggest that according to the need and similar-treatment-for-similar-cases criteria of equity, there have indeed been considerable improvements in the equity of access profile in the U.S. over the past 25 years. These caveats are also borne out by these data:

(1) The poor are much less apt than the nonpoor to see a dentist during a year, (for pregnant women) to visit a doctor during the first trimester of pregnancy, or to have adequate private or public third-party coverage.

(2) The relative gains of the poor in terms of overall physician visit rates may be leveling off or, to some extent, diminishing when compared with the nonpoor.

(3) Persons with low incomes may be using services in response to symptoms about as appropriately as persons with high incomes, but other indicators suggest that relative to experienced disability or chronic illness, complete parity between income groups may still not have been achieved for all age groups.

(4) Pockets of inequity continue to exist, depending on the measure of access that is considered, the type of community in which people live, and the social and demographic characteristics of persons within the community.

† This is a revised version of the ratio shown in Table 4 that is based on a reduced number of the symptoms included in the national study.

Governmental Programs and Equity. Clearly the two major government financing initiatives begun in 1966 that have most influenced the population's use of medical services are Medicare for the elderly (and later the disabled) and Medicaid for those with low incomes. In the later 1970s and 1980, reports of increasing state restrictions on eligibility, benefits, and reimbursements for the Medicaid Program prompt us to examine summary trends in Medicaid and Medicare performance that might change the overall access picture in the United States. Also, changes in various other governmental efforts to provide service to particularly vulnerable groups and those who "fall through the cracks" between private and government financing systems (neighborhood and later community health centers, maternal and child health care, migrant and rural health care) have the potential to influence the access picture. Consequently, we examined trends in the Community Health Centers Program in particular, as well as Medicaid and Medicare.

Table 11 shows steady increases in Medicaid payments from about $3.5 billion in 1968 to over $20 billion in 1979. The number of Medicaid recipients more than doubled from 11.5 million in 1968 to about 23.5 million in 1976. However, in each of the last three years for which we have information the number declined, dropping to a little more than 21 million in 1979. Although we were cautioned that the apparent recent drop in recipients may be influenced by modified reporting in some states,[45] the impact of these changes on the overall summary figures is not clear. Further, the trend the past few years is parallel to the reports in numerous states of fiscal problems in the Medicaid Program with resultant cuts in services.

The annual payments per recipient have grown each year from $300 in 1968 to $950 in 1979. When these payments are adjusted for the increasing price of medical care the increase is not consistent. The constant dollar value of services received in 1979 is actually $424 or about 40% more than in 1968.

Since its inception, Medicare has increased its enrollment each year starting with 19 million in 1966 and reaching almost 27 million in 1978 (Table 12). Contributing to the increase has been the growing number of persons 65 years of age and over as well as the inclusion in the program of other categories of people (those on kidney dialysis and the disabled).

Another Medicare-based indicator of access is the proportion of enrollees who chose to purchase supplementary medical insurance. This proportion has increased each year from 92.4% in 1966 to 97.4% in 1978. The average monthly premium for this supplementary insurance rose from $3.00 in 1966 to about $8.00 in 1978. Adjusting by the Consumer Price Index for all items suggests that recipients in 1978 were paying about one-third more for their medical insurance in 1978 than was the case in 1966.

TABLE 11. Medicaid Payments Adjusted for Increases in Recipients and Prices, Fiscal Years 1968–1978

Fiscal Year	Total (Federal and State) Medicaid Payments (Thousands)*	Yearly Number of Medicaid Recipients (Thousands)†‡	Medical Care Price Index§	Per Cent Growth of Medical Care Costs	Annual Payments Per Medicaid Recipient ($)	Payment Per Recipient in Constant Dollars
1968‖	3,451,376	11,500	100.0	6.1	300	300
1969‖	4,351,486	12,060	106.9	6.9	361	338
1970‖	5,093,901	14,507	113.7	6.4	351	309
1971	6,345,199	17,965	121.0	6.4	353	292
1972	7,346,131	17,990	124.9	3.2	408	327
1973	8,713,761	18,818	129.8	3.9	463	357
1974	9,737,398	20,842	141.9	9.3	467	329
1975	12,086,166	21,197	158.9	12.0	570	359
1976	13,977,348	23,462	174.1	9.6	596	342
1977	16,354,599	22,814	190.8	9.6	717	376
1978	17,965,000	22,198	206.4	8.9	809	392
1979	20,200,000	21,200	224.5	8.7	953	424

Source: Reference 45; Personal Communication: U.S. Department of Health and Human Services, Health Care Financing Agency, Medicaid Bureau, Region V (Chicago, Illinois).

* Source: "State Expenditures for the Medical Assistance Program" except for FY 1978 which comes from the Appendix to the Budget of the U.S. Government for FY 1980. See Technical Notes 1, 2 and 5 in Appendix.

† Includes some recipients of aid under nonfederally matched programs.

‡ Source: "Medicaid State Tables" except for 1971, 1977 and 1978 which are from "Medicaid Statistics." See Technical Notes 3, 4, 5, 6 and 7 in the Appendix.

§ Bureau of Labor Statistics Medical Care Price Index with adjustments to make 1968 = 100.

‖ Includes payments under the Kerr-Mills program.

TABLE 12. Trends in the Medicare Program

Year	Number Enrolled for Hospital Insurance	Per Cent of Hospital Enrolled with Supplemental Medical Insurance	Total Reimbursement for Hospital and Supplemental Medical Insurance	Average Expenditure/ Recipient (Current $)	Average Expenditure/ Recipient (1966 $)	Inpatient Hospital Deductible (Current $)	Inpatient Hospital Deductible (1966 $)	Supplementary Medical Insurance Average Premium/Month (Current $)	Supplementary Medical Insurance Average Premium/Month (1966 $)
1966*	19,082	92.4	886,943	46	46	40	40	3.00	3.00
1968	19,770	95.1	5,289,026	268	236	40	37	4.00	3.73
1970	20,361	96.2	6,605,683	324	251	52	43	4.70	3.93
1972	21,115	96.4	8,165,978	387	273	68	53	5.50	4.27
1973†	23,302	96.5	8,642,782	371	252	68	50	6.05	4.70
1975	24,640	97.0	14,523,573	589	326	92	59	6.70	4.04
1976	25,312	97.2	17,350,766	685	346	104	59	6.95	3.97
1977	26,094	97.2	20,869,055	800	369	124	66	7.45	3.99
1978	26,777	97.4	23,955,521	895	381	144	72	7.95	4.00

Source: References 44, 46–50. * Program in effect half the year. † From 1973 on statistics include disabled persons under 65 years of age.

Total reimbursements for hospital and supplemental medical insurance reached almost $24 billion in 1978. The average reimbursement per enrollee increased from $268 in 1968 (the first full year experience shown in Table 12) to almost $900 in 1978. Adjustment for inflation in the price of medical care suggests that the amount spent per recipient per year in 1966 dollars actually increased from $236 in 1968 to $381 in 1978, an increase of around 60%.

The last indicator of Medicare access we examine is the deductible for hospital services, or the amount the hospitalized patient must pay before Medicare takes over. It was $40 in 1966 and had increased to $144 by 1978. Adjusting for inflation, the increase was from $40 to $72.

The Community Health Centers Program, although much smaller than the overall Medicaid Program in terms of expenditures and users, showed steady increases in the number of people its various projects served—from 1.2 million in 1974 to 4.2 million in 1980 (Table 13). Its distribution of revenues suggests that many of its services and the patients it serves are not covered under Medicare and Medicaid; in 1980 about 16% of its revenues came from Medicare and Medicaid. Despite the increases in users (250%) and encounters (200%) between 1974 and 1980, appropriations in constant 1974 dollars were actually less in 1980 (174 million) than in 1974 (197 million).

Thus, our examination of Medicare and the Community Health Centers Program suggests that coverage in terms of the numbers of people served and services per recipient continued to increase in recent years. Potential access barriers in Medicare might be reflected, however, in the increasing cost of deductibles for hospitalization and the supplementary medical insurance premiums. In the case of the Community Health Centers Program, appropriations do not appear to be keeping pace with the number of services provided, although the Centers are securing revenues from a number of different sources.

The Medicaid picture is different. Although the payment per recipient in constant dollars showed some increase in recent years (about 24% between 1976 and 1979), the number of recipients apparently declined by over two million in this same period. Assuming this decline is not simply a matter of record-keeping, it may represent a trend to increasing access barriers with time. During a similar time period (1976–1978) the number of people with below-poverty-level incomes declined by only one-half million.[50] Given relatively high rates of unemployment, a stagnant economy, and difficulties of local governments in maintaining their own medical care programs, it seems unlikely that other sources can cope with the decline in service delivery and financing resulting from Medicaid cutbacks.

TABLE 13. Community Health Centers

	1974	1976	1978	1980
Appropriations (Current $)	197m	197m	247m	320m
Appropriations (1974 $)	197m	155m	162m	174m
Users	1.2m	2.2m	3.0m	4.2m
Encounters	5.6m	10.0m	13.3m	16.8m
Projects	157	302	591	872

1976 Survey of CHC Users Showed:
71% under poverty level
80% minority
49% have no employed family member
41% under 18

CHC revenues from first and third parties

	1974	1977	1978	1979	1980
Medicare		2.9%	2.3%	2.4%	2.5%
Medicaid		15.1	12.4	12.7	13.3
Title XX		0.3	0.3	0.3	0.3
Private insurance		6.1	6.5	6.7	7.0
Other third party		2.7	3.2	3.3	3.4
Patient fees		9.3	12.5	12.9	13.5
Total (%)	17.4	36.4	37.2	38.3	40.0

Note: In 1974 NCHS began a major management push to improve project collections from first and third parties, and later years show results of these efforts. Clearly, however, Medicaid is *not* providing the major financing of CHCs. In fact, the Medicaid dropoff after 1977 reflects: 1) increased CHC activity in rural areas with lower percentages of Medicaid recipients, and 2) increasing state restrictions on eligibility, benefits, and reimbursements.

Source: Personal Communication: Ronald H. Carlson, Associate Administrator for Planning, Evaluation and Legislation, Health Services Administration, Department of Health and Human Services.

Health Policy and the Current Profile of Equity

Issues to be addressed in any policy deliberations concerning the problems of equity include: (1) are the problems of sufficient magnitude or importance to warrant the expenditure of resources, and (2) if so, what means might be effective in solving them?

Consideration of the first issue must draw upon some of the questions raised in the initial sections of this paper. Inequities do remain that are based upon the need and similar treatment criteria. The magnitude of the problem is less than in the past, but there is no evidence that substantial narrowing of existing gaps is taking place or that current trends will continue should there be significant reductions in programs previously responsible for initiating and sustaining these trends.

Practical questions to consider in deciding whether the problems that remain are important enough to warrant further investment of resources include:

(1) What are the value judgments implied in defining a particular observed "inequity" as a problem?

(2) What criteria should be used to measure how much of a difference signals a "serious" problem, i.e., ± "X"% of the national average?

(3) Will addressing this problem bring about other positive outcomes, i.e., improved health status or more consumer satisfaction?

(4) Are the resources to bring about the desired change available from either the individual or the delivery system?

(5) Will achieving the objective of equity of access mean that other (e.g., cost or quality) goods are compromised?

(6) What are the likely outcomes if no attention is directed to the problem or if existing programs dealing with the problem are discontinued?

These are questions of practical importance in fluid and stressful economic and political times. Assuming that some of the issues identified here are ones that are singled out for attention, we are then called upon to address the second question of the means by which those desired ends can be effected. As mentioned in preceding sections of the paper, the contribution that health services research has made (more or less successfully) in answering this question are efforts to model the "if-then" of various policy choices. There is evidence, for example, that financing initiatives have been responsible primarily for narrowing the overall use differentials by income. On the other hand, health-care reorganization strategies such as community health centers or comprehensive family-centered HMOs also may have contributed to these outcomes. Efforts to focus on solutions to the problems that remain should build on existing empirical work that attempts to estimate the coefficient of association, and/or causality, between certain desired outcomes and mutable policy-relevant means to effect them.

Future empirical work on the equity of access concept could aid the process of decisionmaking just outlined by:

(1) refining further the means of measuring need and use, and the relationship between the two;

(2) developing and applying various criteria of appropriateness;

(3) estimating program impact; and

(4) developing operational indicators of explicit tradeoffs in the cost, quality, and access objectives.

Forces in our society that portend increasing emphasis on cost control at the expense of quality and, in particular, access objectives include:

(1) the current conservative political climate highlighted by the efforts to cut government spending;

(2) the concern about inflation, with medical care singled out as a major contributor;

(3) increasing life expectancy and reduced infant mortality and disease-specific death rates (such as heart disease); and

(4) the growing emphasis on good health habits (diet, exercise, self diagnosis and treatment).

Under these circumstances even the concept of medical care as a right may be increasingly questioned. The Reagan administration has proposed substantial cuts in Medicaid spending. The criteria of social justice dominant in our society with respect to the distribution of health services may shift toward the marketplace, merit, and societal contribution, and therefore away from need and similar treatment.

The shift might be buttressed by the evidence that the poor currently have higher admission rates and more physician visits than the rest of the population; an inability to show that much medical care has a positive impact on traditional health-status measures; continued improvement in these health-status measures (mortality rates); and the increasingly held belief that the individual may be largely responsible for good health status through personal habits or "meritorious actions."

Our findings suggest that the access picture is much more complex than the simple gross hospital and physician use rates indicate. Further, we may be at a critical point as far as slowing or reversing trends in access that have been documented over the past 30 years. The indicators in this report (and it is hoped that better ones will be developed) need to be continually updated and analyzed. This activity will not necessarily alter the criteria currently in vogue for judging equity of access and resulting policy and programs. At least these indicators can relay in a timely fashion to our leaders and the public some of the costs (or benefits) of policy decisions concerning health services delivery and financing.

Acknowledgments

The authors acknowledge the contributions of Ron Wilson (National Center for Health Statistics), Ron Carlson and Ann Zuvekas (Public Health Service) who provided data reported in this paper; of Charles Harris (Methodist Medical Center, St. Louis, MO) for sharing his work on ethics and health policy with us; of Susan Turk and Terry Herold (CHAS) for assembling references and information for the paper; and of Joyce Van Grondelle (CHAS) for typing the manuscript.

References

[1] Blackstone WT. *On health care as a legal right: an exploration of legal and moral grounds.* Georgia Law Review 1976;10:391.

[2] Sidel VW. *The right to health care: an international perspective.* In: Bioethics and Human Rights. Boston: Little, Brown, 1978.

[3] White RL. Right to Health: The Evolution of an Idea. Iowa City: The University of Iowa, 1971.

[4] Carlson R. The End of Medicine. New York: John Wiley and Sons, 1975.

[5] Illich I. Medical Nemesis: The Expropriation of Health. Toronto: Bentam Books, 1976.

[6] Blumstein JF, Zubkoff M. *Public choice in health: problems, politics and perspectives on formulating national health policy.* J Health Polit Policy Law 1979;4:382.

[7] Mechanic D. *Rationing health care: public policy and the medical marketplace.* Hastings Cent Report 1976;6:34.

[8] Sparer EV. *The legal right to health care: public policy and equal access.* Hastings Cent Rep 1976;6:39.

[9] McGarvey MR. *Some considerations regarding ethics and the right to health care.* In: Bioethics and Human Rights. Boston: Little, Brown, 1978.

[10] Gustafson JM. *Interdependence, finitude, and sin: reflections on scarcity.* Journal of Religion 1977;57:156.

[11] Hiatt HH. *Protecting the medical commons: who is responsible?* N Eng J Med 1975;293:235.

[12] Mechanic D. *Ethics, justice, and medical care systems.* Ann AAPSS 1978;437:74.

[13] Weinstein MC, Stason WB. *Allocating resources: the case of hypertension.* Hasting Cent Rep 1977;7:24.

[14] Levine C. *Ethics and health cost containment.* Hastings Cent Rep 1979;9:10.

[15] Navarro V. *Health and the corporate society.* Social Policy 1975;5:41.

[16] Pellegrino Ed. *Medical morality and medical economics.* Hastings Cent Rep 1978;8:8.

[17] Aday LA, Andersen R. *A framework for the study of access to medical care.* Health Serv Res 1974;9:208.

[18] Aday LA, Andersen R. Development of Indices of Access to Medical Care. Ann Arbor MI: Health Administration Press, 1975.

[19] Aday LA, et al. Health Care in the U.S.: Equitable for Whom? Beverly Hills CA: Sage Publications, 1980.

[20] Outka G. *Social justice and equal access to health care.* Perspec Biol Med 1975;18:185.

[21] Navarro V. *The political economy of medical care: an explanation of the composition, nature and functions of present health sector of the United States.* Int J Health Serv 1975;5:65.

[22] Navarro V. *Justice, social policy, and the public's health.* Med Care 1977;15:363.

[23] Fuchs V. *Concepts of health—an economist's perspective.* J Med Philos 1976;1:229.

[24] Garfinkel I. *Equal access, minimum provision, and efficiency in financing medical care.* J Hum Resour 1972;7:242.

[25] Fried C. *Equality and rights in medical care.* Hastings Cent Rep 1976;6:29.

[26] Fein R. *On achieving access and equity in health care.* Milbank Mem Fund Q 1972;50:157.

[27] Bryant JH. *Principles of justice as a basis for conceptualizing a health care system.* INT J HEALTH SERV 1977;7:707.

[28] Harris C. JUSTICE AND VALUE IN HEALTH POLICY. Columbia: University of Missouri, 1980.

[29] Verbrugge LM. EQUITY OF ACCESS TO HEALTH CARE: A RESEARCH REVIEW AND PROSPECTUS. Ann Arbor: Institute for Social Research, University of Michigan, 1980.

[30] Aguirre MN, et al. RESPONSE: THE CONSUMER VIEWPOINT: "WHAT IS HEALTH CARE AND WHAT DO WE WANT?" DHEW Publications No. (HRA) 79-632. Washington, DC: U.S. Government Printing Office, 1980.

[31] Mechanic D. *Correlates of physician utilization: why do major multivariate studies of physician utilization find trivial phychosocial and organization effects?* J HEALTH SOC BEHAV 1979;20:387.

[32] Jonsen AR, Butler LH. *Public ethics and policy making.* HASTINGS CENT REP 1975;5:19.

[33] Sloan FA. *Toward measures of equity in the delivery of health care services.* In: NOURISHING THE HUMANISTIC IN MEDICINE. Pittsburgh: University of Pittsburgh Press, 1979.

[34] Andersen R, et al. TWO DECADES OF HEALTH SERVICES: SOCIAL SURVEY TRENDS IN USE AND EXPENDITURE. Cambridge, MA: Ballinger, 1976.

[35] National Center for Health Statistics. HEALTH: UNITED STATES, 1980. DHHS Publication No. (PHS) 81-1232. Washington, DC: U.S. Government Printing Office, 1980.

[36] Aday LA, et al. *Overview of a design to evaluate the impact of community hospital-sponsored primary care group practices.* MEDICAL GROUP MANAGEMENT 1978;256:42.

[37] Block JA, et al. *Physicians and hospitals: providing primary care.* MEDICAL GROUP MANAGEMENT 1978;25:35.

[38] Andersen R, Anderson OW. A DECADE OF HEALTH SERVICES: SOCIAL SURVEY TRENDS IN USE AND EXPENDITURE. Chicago: The University of Chicago Press, 1967.

[39] Andersen R, Anderson OW. *Trends in the use of health services.* In: HANDBOOK OF MEDICAL SOCIOLOGY. Englewood Cliffs: Prentice-Hall, 1979.

[40] National Center for Health Statistics. *Persons hospitalized by number of hospital episodes and days in a year: United States—July 1965-June 1966.* Series 10, No. 50. Washington, D.C.: U.S. Government Printing Office, 1969.

[41] Phelps CE. *Effects of insurance on demand for medical care.* In: EQUITY IN HEALTH SERVICES: EMPIRICAL ANALYSES IN SOCIAL POLICY. Cambridge, MA: Ballinger, 1975.

[42] Newacheck PW, et al. *Income and illness.* MED CARE 1980;18:1165.

[43] National Center for Health Statistics. *Health characteristics of persons with chronic activity limitations: United States—1974.* Series 10, No. 112. Washington, DC: U.S. Government Printing Office, 1976.

[44] National Center for Health Statistics. HEALTH: UNITED STATES, 1979. DHEW Publication No. (PHS) 80-1232. Washington, DC: U.S. Government Printing Office, 1980.

[45] Health Care Financing Administration. *Data on the Medicaid Program: Eligibility, Services and Expenditures.* Baltimore: Medicaid/ Medicare Management Institute, 1979.

[46] Social Security Administration. SOCIAL SECURITY BULLETIN, ANNUAL STATISTICAL SUPPLEMENT, 1975. Washington, DC: U.S. Government Printing Office, 1975.

[47] Social Security Administration. SOCIAL SECURITY BULLETIN, ANNUAL STATISTICAL SUPPLEMENT, 1977–79. Washington, DC: U.S. Government Printing Office, 1979.

[48] U.S. Department of Commerce. STATISTICAL ABSTRACT OF THE UNITED STATES, 1978. Washington, DC: U.S. Government Printing Office, 1978.

[49] Waldo DR. HEALTH CARE FINANCING TRENDS 1980–2. Washington, D.C.: Office of Research, Demonstrations and Statistics, HCFA.

[50] U.S. Department of Commerce. *Population Profile of the United States, 1979.* Current Population Reports. Series P-20, No. 350. Washington, DC: U.S. Government Printing Office, 1980.

Uninsured and Underserved: Inequities in Health Care in the U.S.

Karen Davis, Ph.D.*
Diane Rowland, M.P.A.†

The United States has one of the highest-quality, most sophisticated systems of medical care in the world. Most Americans take for granted their access to this system of care. In times of emergency or illness, they can call upon a vast array of health resources—from a family physicians to a complex teaching hospital—assured that they will be welcome and that care will be affordable, with health insurance coverage to pick up the tab for the majority of bills incurred.

For a surprisingly large segment of the U.S. population, however, this ease of access to care does not exist. At any point in time, over 25 million Americans have no health insurance coverage—from either private health insurance plans or public programs (Kasper *et al.*, 1978). Without health insurance coverage or ready cash, such individuals can be and are turned away from hospitals even in emergency situations (Committee on Energy and Commerce, 1981). Some neglect obtaining preventive or early care, often postponing care until conditions have become life-threatening. Others struggle with burdensome medical bills. Many come to rely upon crowded, understaffed public hospitals as the only source of reliable, available care.

The absence of universal health insurance coverage creates serious strains in our society. These strains are felt most acutely by the uninsured poor, who must worry about family members—a sick child, an adult afflicted with a deteriorating chronic health condition, or a pregnant mother—going without needed medical assistance. It strains our image as a just and humane society when significant portions of the population

*Professor and †Research Associate, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, Maryland.

September 1982

endure avoidable pain, suffering, and even death, because of an inability to pay for health care. Those physicians, other health professionals, and institutions that try to assist this uninsured group also incur serious strain: Demands typically far outstrip available time and resources. Strain is also felt by local governments whose communities include many uninsured persons, because public hospitals and death centers inevitably incur major financial deficits.

As serious as these strains have been in the last five years, the years ahead promise to strain the fabric of our social life even more seriously. Major reductions in funding for health services for the poor have been made in the last year; further reductions are likely. Deepening economic recession, high unemployment, and declining sales revenues are strapping the fiscal resources of state and local governments. Their ability to offset Federal cutbacks seems limited. Nor can the private sector be expected to bridge this gap. The health industry is increasingly becoming an entrepreneurial business endeavor—with little room for charitable actions.

It is especially timely, therefore, to review what we know about the consequences of inadequate health insurance coverage for certain segments of our population. The first section of this paper presents information on the number and characteristics of the uninsured, while the second section describes patterns of health care utilization by the uninsured. The third section describes the policy implications of these facts, and this paper concludes with recommendations for future public policy to ensure access to health care for all.

Who Are the Uninsured?

The 1977 National Medical Care Expenditure Survey (NMCES) provides extensive information on the health insurance coverage of the U.S. population. Six household interviews of a nationwide sample of over 40,000 individuals were conducted over an 18-month period during 1977 and 1978. By following the interviewed population for an entire year, NMCES provided a comprehensive portrait of health insurance coverage, including changes in health insurance status during the course of that year.

In the NMCES results, individuals classified as insured are those who were covered throughout the year by Medicaid, Medicare, Blue Cross/Blue Shield, or commercial health insurance, or who were enrolled in a health maintenance organization. Differences in scope of coverage among the insured were not currently available. Therefore, it may be that many individuals in the insured category actually had very limited health insurance coverage, leaving them basically uninsured for most services. The uninsured fall into two groups: the always unin-

Table 1: Insurance Status During Year, by Age and Race (1977)

	Total	Always Uninsured	Uninsured Part of Year	Always Insured
	Number (millions)			
Persons under 65	189.8	18.0	15.8	156.0
White	163.7	14.5	12.5	136.7
Black and other	26.1	3.5	3.3	19.3
Persons 65 and over	22.3	0.1	0.1	22.1
White	20.2	0.07	0.09	20.0
Black and other	2.1	0.03	0.01	2.1
TOTAL	212.1	18.1	15.9	178.1
	Percent (%)			
Persons under 65	100	9.5	8.3	82.2
White	100	8.9	7.6	83.5
Black and other	100	13.3	12.7	74.0
Persons 65 and over	100	0.4	0.5	99.1
White	100	0.3	0.5	99.2
Black and other	100	1.0	0.8	98.2
TOTAL	100	8.6	7.5	83.9

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

sured and the sometimes uninsured. The always uninsured are individuals without Medicare, Medicaid, or private insurance coverage for the entire year. The sometimes uninsured are those who were covered by public or private insurance part of the year but were uninsured the remainder of the year.

A snapshot view of the uninsured at a given point in time understates the number of people who spend some portion of the year uninsured; at any one time, there are over 25 million uninsured Americans, but as many as 34 million may be uninsured for some period of time during the year. Approximately 18 million are without insurance for the entire year, and 16 million are uninsured for some portion of the year (Wilensky and Walden, 1981). Changes in insurance status during the year are generally the result of loss of employment, change in employment, change in income or family situation that alters eligibility for Medicaid, or loss of private insurance when an older spouse retires and becomes eligible for Medicare.

The 34 million uninsured are persons of all incomes, racial and ethnic backgrounds, occupations, and geographic locations. In some cases whole families are uninsured, while in others coverage is mixed depending on employment status and eligibility for public programs (Kasper *et al.*, 1978). However, the poor, minorities, young adults, and rural residents are more likely than others to be uninsured. As noted in Table 1, over one-quarter of all blacks and minorities are uninsured during the year—a rate 1½ times that of whites. This disparity holds across the demographic and social characteristics of the uninsured (Wilensky and Walden, 1981).

Age. The uninsured population, whether covered for all or part of a year, is almost entirely under age 65. Nearly one-fifth of the nonelderly population is uninsured for some or all of the year. As shown in Table 1, less than 1% of the aged, barely 200,000 persons, are uninsured during the year. This is primarily attributable to Medicare, which provides basic coverage for hospital and physician services to most older Americans. The success of Medicare in providing financial access to health care for the elderly is demonstrated by the extensive coverage of the elderly today in contrast to the dramatic lack of insurance prior to implementation of Medicare in 1966 (Davis, 1982). Medicaid and private insurance help to fill the gap for those elderly persons ineligible for Medicare because they lack sufficient quarters of Social Security earnings contributions. The uninsured elderly are primarily individuals with incomes above the eligibility levels for welfare assistance and Medicaid.

Examination of the uninsured by age group reveals that young adults are the group most likely to be uninsured. As highlighted in Table 2, almost one-third of all persons aged 19-24 are uninsured during the course of a year. Roughly 16% of this age group are without coverage all year, and an additional

14% lack coverage at least part of the year. This rate is nearly double that of other age groups. A variety of factors undoubtedly contribute to this situation. Young adults frequently lose coverage under their parents' policies at age 18. Many young adults may elect to forego coverage when it is available, since coverage is costly and they assume themselves to be relatively healthy. High youth unemployment as well as employment in marginal jobs without health benefits make insurance difficult to obtain or afford for this group.

Table 2: Percentage Insured During Year, by Selected Population Characteristics (1977)

	Uninsured During Year (%)	Always Uninsured (%)	Uninsured Part of Year (%)
Age			
Under 65	17.8	9.5	8.3
Under 6 years	19.6	8.3	11.3
6 to 18 years	16.1	8.6	7.5
19 to 24 years	30.3	16.0	14.3
25 to 54 years	16.1	8.7	7.4
55 to 64 years	12.6	8.2	4.4
Age 65 and over	0.9	0.4	0.5
Occupation			
Farm	22.3	15.9	6.4
Blue-collar	19.8	11.3	8.5
Services	20.8	11.9	8.9
White-collar	12.6	5.6	7.0
Region			
Northeast	10.7	5.4	5.3
North Central	12.5	5.7	6.8
South	20.5	11.6	8.9
West	20.8	11.7	9.1
TOTAL	16.1	8.6	7.5

Source: Wilensky and Walden, *Minorities, Poverty, and the Uninsured*, and data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

Employment. Employment status and occupation are important factors in assessing the likelihood of being uninsured for all or part of a year. Most U.S. workers receive their health care coverage through the workplace. But insurance coverage varies widely depending on the characteristics of the employer (Taylor and Lawson, 1981). Employees of small firms are less likely to be insured than employees of large firms. For example,

45% of employees in firms of 25 or fewer employees do not have employer health insurance, compared with only 1% in firms with more than 1000 employees. Yet small firms employ over 20% of all workers. Unionized firms are six times more likely to have employee health insurance than are nonunionized firms (Davis, 1975).

As Table 2 shows, insurance status varies by type of employment. Nearly one-quarter of all agricultural workers are uninsured during the year, with 16% uninsured for the entire year. As expected, white-collar workers are the most likely to be insured, while blue-collar and service workers fare only somewhat better than agricultural workers (Wilensky and Walden, 1981). Among blue-collar and service workers, insurance coverage is low in the construction trade, wholesale and retail trades, and service industries, and high in manufacturing. Of manufacturing employees, 96% have health insurance through their place of employment (Davis, 1975).

Residence. These trends in coverage by employment are reflected in the regional picture of insurance status. In the heavily industrial and unionized Northeast and North Central regions of the country, the percentage of uninsured during the year is half that of the South and the West. In these areas where agricultural interests are strong and unionization less extensive, over 20% of the population is uninsured during the course of a year. Of those living in the South and West, 11% are uninsured throughout the year compared with 5% in the Northeast and North Central regions. Similarly, people in metropolitan areas are more likely to be insured than people living outside metropolitan areas (Wilensky and Walden, 1981).

Income and Race. However, while nature of employment and unionization may explain some of the regional variations, a critical underlying factor in the analysis is the distribution of the poverty and minority population. Residents of the South comprise 32% of the total population under age 65. Yet 48% of the nation's minorities live in the South (DHHS, 1982a). The higher concentrations of poor and minority persons in the South in comparison with other parts of the country help explain the high level of uninsured individuals.

Poverty and lack of insurance are strongly correlated. Of poor families with incomes below 125% of the poverty line, 27% are uninsured. The near-poor, with incomes between 125-200% of poverty ($8000-$10,000 for a family of four in 1977), fare only slightly better—with 21% uninsured during the year. As Table 3 shows, the poor are always more likely to be uninsured than the middle and upper-income groups (Wilensky and Walden, 1981).

The limited health insurance coverage for the poor and near-poor demonstrates the limits of coverage under Medicaid. Many assume that Medicaid finances health care services for all

Table 3: Percentage Insured During Year, by Ethnic/Racial Background and Income (1977)

	Uninsured During Year (%)	Always Uninsured (%)	Uninsured Part of Year (%)
White, all incomes	14.0	7.0	7.0
Poor	27.1	13.5	13.6
Other low-income	21.0	10.9	10.1
Middle-income	12.6	6.3	6.3
High-income	8.8	4.2	4.6
Black, all incomes	23.2	9.7	13.5
Poor	32.2	10.6	21.6
Other low-income	26.6	11.9	14.7
Middle-income	17.4	8.6	8.8
High-income	12.4	7.1	5.3
Hispanic, all incomes	24.3	12.8	11.5
Poor	29.6	9.5	20.1
Other low-income	32.0	18.2	13.8
Middle-income	17.7	12.4	5.3
High-income	20.0	12.3	8.0

Source: Wilensky and Walden, *Minorities, Poverty, and the Uninsured.*

of the poor. However, many poor persons are ineligible for Medicaid due to categorical requirements for program eligibility and variations in state eligibility policies. Two-parent families are generally ineligible for Medicaid, and single adults are covered only if they are aged or disabled (Davis and Schoen, 1978). Moreover, many states have established income eligibility cutoffs well below the poverty level. Many states have not adjusted income levels to account for inflation, resulting in a reduction in the number of individuals covered over the last few years (Rowland and Gaus, 1981). As a result of the restrictions on Medicaid coverage, about 60% of the poor are not covered by Medicaid. This situation can only be expected to worsen as the recession swells the numbers of poor and near-poor, while cutbacks in social programs and Medicaid further erode the health coverage now available to some of the poor.

Thus, while the poor are obviously the least able to pay for care directly, they are the most likely to be without either Medicaid or private insurance. The poor are twice as likely to be uninsured as the middle class and three times as likely as those in upper-income groups. Lack of insurance is inversely related to ability to bear the economic consequences of ill health.

Blacks, Hispanics, and other minorities are also more likely to be uninsured than whites, regardless of income group; and poor blacks are the most likely to be uninsured. (See Table 3.) Nearly one-third of poor blacks are uninsured during a year. If you are poor and a member of a minority group, your chances of being uninsured are four times as great as for a high-income white.

Yet this relationship between race and income shown in Table 3 actually understates the situation, because the aged are included in the population analyzed. The aged are overrepresented in the lower-income groups, but as Table 1 shows, almost all of the aged are insured. Thus, inclusion of the aged in Table 3 tends to overstate the insured status of the nonelderly poor.

In Table 4, regional and racial differences in insurance coverage are shown for the population under age 65. When the aged are excluded from the analysis, the differentials become even more striking. Southerners are nearly 1½ times as likely to be uninsured as those from other parts of the country. But blacks in the South are 1½ times more likely to be uninsured as are whites from the South or non-Southern blacks. Southern blacks are twice as likely to be uninsured as non-Southern whites.

Similarly, when differences in insurance status are assessed from the perspective of metropolitan versus nonmetropolitan areas, blacks fare much worse than whites. Over 16% of nonelderly residents of Standard Metropolitan Statistical Areas (SMSAs) are uninsured compared with over 21% of those residing in non-SMSA areas. But, for minorities living outside SMSAs, almost 40% are uninsured—a rate twice that of whites residing in non-SMSA areas and 2½ times that of whites in SMSAs.

Thus, health insurance coverage in the U.S. is to some extent a matter of luck. Those fortunate enough to be employed by large, unionized, manufacturing firms are also likely to be fortunate enough to have good health insurance coverage. Those who are poor, those who live in the South or in rural areas, and those who are black or minority group members are more likely to bear the personal and economic effects of lack of insurance and the consequent financial barriers to health care access.

Utilization of Health Services by the Uninsured

With the investment in primary care made by Federal programs in the late 1960s and 1970s, significant progress was achieved in improving access to primary care for the poor and other disadvantaged groups. Numerous studies have been con-

Table 4: Percentage of Persons Under Age 65 Uninsured During Year, by Race and Residence (1977)

	Population (millions)	Uninsured During Year (%)	Always Uninsured (%)	Uninsured Part of Year (%)
South	60.5	22.4	12.7	9.7
White	47.9	20.4	11.8	8.6
Black and other	12.6	30.0	16.2	13.8
Non-South	129.3	15.7	8.0	7.7
White	115.8	14.9	7.7	7.2
Black and other	13.5	22.2	10.7	11.5
SMSA	132.6	16.3	8.2	8.1
White	111.3	14.9	7.6	7.3
Black and other	21.3	23.2	11.1	12.1
Non-SMSA	57.2	21.4	12.5	8.9
White	52.5	19.9	11.6	8.3
Black and other	4.7	38.2	23.3	14.9
TOTAL	189.8	17.8	9.5	8.3

ducted to examine trends in access to health care (Davis *et al.*, 1981). Virtually all studies conclude that income-related differentials in utilization of physician services and preventive services have narrowed.

In the early 1960s the nonpoor visited physicians 23% more frequently than the poor, even though the poor, then as now, were considerably sicker than the nonpoor. By the 1970s the poor visited physicians more frequently than the nonpoor, and more in accordance with their greater need for health care services. Blacks and other minorities also made substanial gains over this period. Utilization of services by rural residents also increased relative to urban residents (Davis and Schoen, 1978).

Use of preventive services by the poor, minorities, and rural residents continues to lag well behind use by those not facing similar barriers to health care. Some studies have also found that these differentials continue to exist for all disadvantaged groups even when adjusted for the greater health needs of the disadvantaged (Davis *et al*, 1981).

The major difficulty with past studies, however, is that they have not examined insurance coverage of subgroups of the poor to detect the cumulative impact of lack of financial and physical access to care. How do uninsured blacks in rural areas fare in obtaining ambulatory care services? Can nearly all disadvantaged persons get care from public hospitals or clinics, or do those facing multiple barriers to care simply do without?

Ambulatory Care. New data from the 1977 National Medical Care Expenditure Survey (NMCES) shed some light on the cumulative effect of multiple barriers to care. Table 5 presents data on use of physicians' services from NMCES for the population under age 65. Insured persons are those covered during the entire year; the uninsured are those uninsured for the entire year. Those insured for part of the year are excluded; presumably their utilization resembles that of the insured for the portion of the year in which they are insured and that of the uninsured for the portion of the year in which they are uninsured.

Most striking is the extent to which insurance coverage affects use of ambulatory care. As shown, the insured average 3.7 visits to physicians during the year, compared with 2.4 visits for the uninsured. That is, the insured receive 54% more ambulatory care from physicians than do the uninsured. It should be noted that some of the insured may only have insurance coverage for inpatient hospital care, not ambulatory care. Better data on ambulatory care insurance coverage might indicate even greater differentials in use of ambulatory care.

These differentials are even greater for subgroups of the uninsured population. Residence and race further affect utilization of ambulatory services. The lowest utilization of ambula-

Table 5: Physician Visits per Person Under Age 65, by Insurance Status, Residence, and Race (1977)

	Uninsured	Insured	Ratio
South	2.1	3.5	1.67
White	2.3	3.7	1.61
Black and other	1.5	2.8	1.87
Non-South	2.6	3.8	1.46
White	2.7	3.8	1.41
Black and other	1.9	3.5	1.84
SMSA	2.4	3.8	1.58
White	2.6	3.9	1.50
Black and other	1.7	3.2	1.88
Non-SMSA	2.3	3.3	1.43
White	2.4	3.4	1.42
Black and other	1.6	2.9	1.81
TOTAL	2.4	3.7	1.54

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

tory care occurs for uninsured blacks and other minorities, including Hispanics. These persons use ambulatory services far less than more advantaged groups. For example, uninsured blacks and other minorities in the South make 1.5 physician visits per person annually, compared with 3.7 physician visits for insured whites in the South. That is, to be advantaged multiply leads to a utilization rate almost 2.5 times that of individuals who are disadvantaged multiply.

These data point to the importance of financial and physical barriers to access. It is not the case that the uninsured manage to obtain ambulatory care comparable in amount to that obtained by the insured by relying on public clinics, teaching hospital outpatient clinics, nonprofit health centers, or the charity of private physicians. Without insurance, many simply do without care.

The patterns of utilization for different groups provide some insight into the relative importance of financial, physical, and racial barriers to care. Financial access to care is clearly the most important factor affecting use. Insurance coverage reduces much but not all of the differential in use of ambulatory services. Insured blacks in the South, for example, average 2.8 physician visits annually, compared with 3.5 for insured whites in the South. That is, whites average about 25% more ambula-

tory care than blacks and other minorities, even if both are insured. But this differential is substantially smaller than the 2½ times greater use of physicians between insured Southern whites and uninsured Southern blacks.

Location remains an important determinant of use of physician services. Lack of insurance coverage is more predominant in rural areas; however, even among the insured, urban residents are more likely to receive ambulatory care than are rural residents, whether white or black. (See Table 5.) Among insured groups, rural whites receive 3.4 physician visits annually, compared with 3.9 visits for urban whites. Rural blacks and other minorities with insurance make 2.9 physician visits, compared with 3.2 visits for their insured counterparts in urban areas. That is, a 10-15% differential in use between urban and rural areas occurs even when financial access to care is not a problem. (It should be noted, however, that the quality of insurance for ambulatory care may not be as good in rural areas as in urban areas.)

Racial differentials in utilization of ambulatory care are also ameliorated with insurance coverage. Insurance is particularly helpful in improving access to care for minorities. Insured minorities receive 80-90% more ambulatory care than do uninsured minorities in both rural and urban areas. But even with insurance, strong racial differences persist. For example, as shown in Table 5, insured whites in the South receive almost 30% more care than do insured blacks and other minorities in the South (3.7 visits for insured whites versus 2.8 annual visits for insured blacks and other minorities).

Hospital Care. Despite the common perception that all disadvantaged persons can obtain hospital care from some charity facility, tremendous differentials in use of hospital care also exist by insurance status, residence, and race. As shown in Table 6, the insured receive 90% more hospital care than do the uninsured. Differentials by insurance status are particularly marked in the South and in rural areas. In the South insured persons receive three times as many days of hospital care annually as uninsured persons, regardless of race or ethnicity.

These hospital utilization differentials clearly demonstrate that the insured fare much better than the uninsured in obtaining health care services. Since those with insurance are likely to have basic coverage for hospitalization, the hospital utilization data provide a more accurate assessment of the role of insurance coverage in the use of health care services than do the ambulatory care differentials.

These differentials are breathtaking, and they shatter any complacency about the accessibility of inpatient care. More disaggregated information is essential on the types of conditions for which the insured receive inpatient care and the uninsured do not. Standards for appropriate utilization of hospital

Table 6: Hospital Patient Days per 100 Persons Under Age 65, by Insurance Status, Residence, and Race (1977)

	Uninsured	Insured	Ratio
South	35	104	2.97
White	33	100	3.03
Black and other	40	119	2.98
Non-South	56	84	1.50
White	51	81	1.59
Black and other	89	114	1.28
SMSA	50	86	1.72
White	44	83	1.89
Black and other	70	106	1.51
Non-SMSA	42	99	2.36
White	43	94	2.19
Black and other	39	175	4.49
TOTAL	47	90	1.91

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

services are still the subject of wide debate. Some of the differential between the insured and uninsured seen here may be the result of overutilization of hospital services by the insured. However, this is unlikely to explain the entire differential.

Some of the greater utilization of hospital care by the insured may represent self-selection. Those who expect to be hospitalized may obtain such coverage. Hospitalization may itself result in Medicaid coverage of some of the poor and near-poor. The spend-down provision of the Medicaid program enables some low-income individuals who would otherwise be ineligible for Medicaid to receive coverage when they incur large medical bills. In the 29 states that extend Medicaid coverage to the medically needy, an individual whose income is above the Medicaid standard is covered by Medicaid when total income minus the cost of care results in a reduced available income that is below the Medicaid standard. Many of the near-poor could thus become eligible for Medicaid coverage with an expensive hospital episode. Thus, though uninsured throughout most of the year, such individuals would be insured for that part of the year in which they used hospital care. It should be noted, however, that this coverage option is only available in 29 states—and in those states only for individuals meeting the Medicaid categorical requirements (aged, blind, disabled, or family with dependent children). Most two-parent families and

virtually all single individuals and childless couples under age 65 remain ineligible for Medicaid regardless of the size of their medical bills (Rowland and Gaus, 1981).

But these explanations are unlikely to account for a threefold differential in use. Barriers to access to hospital services for the uninsured need to be explored. To what extent do hospitals require preadmission deposits for the uninsured? What are the consequences of such policies on access to care? What hospitals serve the uninsured? Do the uninsured have to travel sizeable distances to obtain services? To what extent do any and all of these factors influence the use of hospital care by the uninsured? Further exploration is certainly warranted.

Health Status and Use of Services. Lower utilization of ambulatory and inpatient care by the uninsured is not a reflection of lower need for health care services. Instead, as measured by self-assessment of health status, the uninsured tend to be somewhat sicker than the insured. As shown in Table 7, 15% of the uninsured under age 65 rate their health as fair or poor, compared with 11% of the insured. Blacks and other minorities in the South systematically rate their health the worst. Of uninsured blacks and other minorities in the South, 19% assess their health as fair or poor, compared with 9% of insured whites outside the South.

Table 7: Percentage of Persons Under Age 65 in Fair or Poor Health, by Insurance Status, Residence, and Race (1977)

	Uninsured (%)	Insured (%)	Ratio
South	18	14	1.30
White	18	13	1.40
Black and other	19	19	1.00
Non-South	12	10	1.24
White	12	9	1.35
Black and other	12	17	.71
SMSA	13	11	1.21
White	12	10	1.29
Black and other	15	18	.88
Non-SMSA	17	12	1.45
White	17	11	1.50
Black and other	17	18	.94
TOTAL	15	11	1.33

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

Table 8: Physician Visits per Person Under Age 65 in Fair or Poor Health, by Insurance Status, Residence, and Race (1977)

	Uninsured	Insured	Ratio
South	3.8	6.1	1.61
White	4.4	6.4	1.45
Black and other	2.2	5.0	2.27
Non-South	4.5	7.4	1.64
White	4.6	7.6	1.65
Black and other	3.5	6.5	1.86
SMSA	4.1	7.2	1.76
White	4.7	7.6	1.62
Black and other	2.3	5.9	2.57
Non-SMSA	4.2	6.3	1.50
White	4.3	6.4	1.49
Black and other	3.2	5.4	1.69
TOTAL	4.1	6.9	1.68

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

One possible explanation of the higher rate of poor/fair health among the uninsured is that the lack of insurance is itself related to health status. Those who rate their health as poor or fair are more likely to be unable to work because of illness than those who rate their health good or excellent. Since insurance coverage in the U.S. is related to employment, those who are unemployed due to poor health are also likely to be without insurance, while the working population enjoys both good health and insurance coverage.

The sick who are uninsured use medical care services less than their insured counterparts. Utilization of ambulatory services adjusted for health status shows that the insured in poor health see a physician 3 times for every 1 visit by the uninsured in poor health. As shown in Table 8, physician visits per person under age 65 in fair or poor health average 6.9 among the insured, compared with 4.1 visits for the uninsured with similar health problems. Blacks and other minorities with fair or poor health who are insured receive twice as much care as their uninsured counterparts.

Among the uninsured in poor or fair health, the differentials in physician visits by race and residence are especially noteworthy. Uninsured whites have greater access to physician services than do uninsured minorities. A Southern white in fair

or poor health sees a physician twice as often as a Southern minority person in fair or poor health. The same relationship exists for utilization of physician services in metropolitan areas. However, the utilization differential between whites and minorities narrows in areas outside the South and in non-metropolitan areas.

The number of physician visits by the uninsured versus the insured in fair or poor health warrants further examination. It is expected that an individual in fair or poor health would require frequent physician visits for diagnosis and treatment of the condition. The average of five to seven visits annually by the insured would appear to provide a reasonable level of physician contact. But, for uninsured minorities in the South in fair or poor health, the average number of visits is two per year. This rate would provide no more than an initial visit and one follow-up visit, which might be insufficient to treat serious or complex illnesses. Thus, lower rates of physician visits could impair adequate treatment and follow-up to promote a rapid recovery.

Dental Care. Dental care, unlike hospital care and most physician services, is not covered under most insurance plans. Therefore, differentials in dental visits between the insured and uninsured are not meaningful. However, the NMCES data do show a striking contrast between dental visits by whites and minorities.

Whites obtain dental care twice as often as minorities, averaging 1.5 visits per year compared to 0.7 visits for minorities. Non-Southern whites had two times the number of visits as non-Southern minorities and over three times the number of visits as Southern minorities. Rural minorities appear to have the least access to dental services.

The significant differential between access to dental services for minorities and whites warrants further examination. The extent to which this differential reflects differences in health practices and attitudes toward dental care or differences in availability and accessibility to dental care should be explored.

Usual Source of Care. The NMCES data confirm other studies that have found disadvantaged groups less likely to have a usual source of ambulatory care and more likely to receive their care from a hospital outpatient department or a clinic than from a physician's office. Table 9, for example, illustrates that 84% of the insured have a physician's office as their usual source of care, compared with 67% of the uninsured. About 50% of uninsured blacks and other minorities have a physician's office as their usual source of care. While this percentage is quite low in comparison with other groups, it does not fit the stereotype that all minorities in urban areas receive the bulk of their care from public facilities or hospital outpatient departments.

Table 9: Percentage of Persons Under Age 65 Whose Usual Source of Care is a Physician's Office, by Insurance Status, Residence, and Race (1977)

	Uninsured (%)	Insured (%)	Ratio
South	66	81	1.22
White	70	82	1.16
Black and other	53	76	1.41
Non-South	68	85	1.25
White	70	86	1.22
Black and other	45	69	1.53
SMSA	63	82	1.31
White	66	84	1.27
Black and other	49	71	1.43
Non-SMSA	73	86	1.19
White	76	87	1.15
Black and other	52	79	1.53
TOTAL	67	84	1.25

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

Uninsured residents of nonmetropolitan areas are more likely to have a physician as a usual soure of care than are residents of a metropolitan area. In nonmetropolitan areas, 73% of the uninsured have a physician as a usual source of care in contrast to only 63% of the uninsured in metropolitan areas. However, nonmetropolitan residents are still likely to have fewer physician visits than their metropolitan counterparts, as shown in Table 5. The nonmetropolitan uninsured get more of their care from physicians, but receive less total care. These differences in utilization among the uninsured undoubtedly reflect differences between metropolitan and nonmetropolitan areas in the availability of alternatives to physician care. Residents of metropolitan areas are more likely to have access to clinics and outpatient hospital services that can substitute for care in physicians' offices.

The metropolitan/nonmetropolitan differential for physician as usual source of care is markedly reduced among the insured. As seen in Table 9, 86% of insured nonmetropolitan residents and 82% of insured metropolitan residents have a physician as a usual source of care. Insurance coverage significantly increases the proportion of minorities who have a physician's office as their usual source of care. Among the minority

uninsured, 49% of those living in metropolitan areas and 52% of those in nonmetropolitan areas have a physician as a usual source of care. In contrast, for insured minorities, 71% in metropolitan areas and 79% outside of metropolitan areas have physicians as a usual source of care. This would suggest that Medicaid coverage enables a substantial number of minorities to obtain care in a physician's office.

Convenience of Care. When they are able to obtain care, the uninsured must travel longer distances than the insured to obtain it. As shown in Table 10, 25% of the uninsured travel 30 minutes or more to obtain care, compared with 18% of the insured. Differentials in travel time between the insured and uninsured are somewhat more marked in rural areas than in urban areas, but travel time is a problem for uninsured persons everywhere. These data suggest not only that the uninsured receive less care, but also that when they do obtain care they do so by searching over a longer distance for providers willing to see them. The effort involved in such a search for care may discourage the use of preventive services, and result in the uninsured only seeking care for serious illness or in crises. This would help explain the lower utilization levels of the uninsured.

Table 10: Percentage of Persons Under Age 65 Traveling More Than 29 Minutes to Receive Medical Care, by Insurance Status, Residence, and Race (1977)

	Uninsured (%)	Insured (%)	Ratio
South	29	21	1.39
White	30	20	1.48
Black and other	28	26	1.09
Non-South	21	16	1.29
White	22	16	1.35
Black and other	17	21	.81
SMSA	22	17	1.27
White	21	16	1.32
Black and other	24	24	1.00
Non-SMSA	29	20	1.46
White	30	20	1.50
Black and other	23	19	1.24
TOTAL	25	18	1.39

Source: Data from the U.S. Department of Health and Human Services, National Center for Health Services Research, National Medical Care Expenditure Survey.

When the uninsured arrive at a care provider, they generally have to wait longer for care to be delivered. Minorities experience longer waiting times than whites. Regardless of residence, the waiting time for insured blacks and other minorities is longer than the waiting time experienced by uninsured whites. Waiting times are longer in the South than in the non-South. Uninsured Southern minority persons experience the longest waiting times; the NMCES data show that they wait one-third longer than do insured Southern whites (DHHS, 1982a).

Policy Implications

The utilization differentials between the insured and uninsured underscore the importance of financial barriers to health care. Lack of insurance coverage is the major barrier. It markedly affects the amount of both ambulatory and inpatient care received. Without insurance coverage, many individuals obviously do without care. Those able to obtain care incur substantial travel and waiting time.

Lack of insurance coverage has three major consequences: It contributes to unnecessary pain, suffering, disability, and even death among the uninsured; it places a financial burden on those uninsured who struggle to pay burdensome medical bills; and it places a financial strain on hospitals, physicians, and other health care providers who attempt to provide care to the uninsured.

Research is limited on both the health of the uninsured and the health consequences of no insurance. Extensive data on utilization patterns by the uninsured disaggregated by residence and race are presented for virtually the first time in this report. But a number of recent studies have shown that medical care utilization has a dramatic impact on health. A recent Urban Institute report by Hadley (1982) explores the relationship between medical care utilization and mortality rates. It contains persuasive evidence that utilization of medical care services leads to a marked reduction in mortality rates. A recent study by Grossman and Goldman (1981) at the National Bureau of Economic Research has found that infant mortality rates have dropped significantly in communities served by federally funded community health centers. This growing body of evidence provides considerable support to the importance of medical care utilization in assuring a healthy population—and at least indirectly provides a basis for concern that the lower rate of medical care utilization by the uninsured contributes to unnecessary deaths and lowered health status.

Lack of insurance coverage also imposes serious financial burdens on those who try to make regular payments to retire enormous debts incurred in obtaining medical care. With the average cost of a hospital stay in the U.S. now in excess of

$2000, few individuals can afford to build payments for hospital care into their monthly living allowance (DHHS, 1982b). Yet, since the uninsured are more likely to be poor, the economic consequences of lack of insurance fall heaviest on those least able to bear the burden.

In addition to its consequences for the uninsured themselves, lack of insurance also takes its toll on the health care system. One result is that the financial stability of hospitals and ambulatory care providers willing to provide charity care for those unable to pay is jeopardized. Health care providers serving the uninsured, particularly inner-city community and teaching hospitals, county and municipal clinics, and community health centers, absorb much of the cost of this care as charity care or a bad debt. Yet this burden is not evenly distributed among hospitals and other providers. A recent study by the Urban Institute found that one-seventh of the national sample of hospitals studied provided over 40% of the free care (Brazda, 1982).

Recent policy measures are likely to exacerbate this situation. The Omnibus Budget Reconciliation Act of 1981 reduced Federal financial participation in Medicaid and curtailed eligibility under the Aid to Families with Dependent Children (AFDC) program. Actions by state governments in response to this legislation could swell the ranks of the uninsured poor by over one million people. Coupled with the highest rate of unemployment since the Great Depression and the loss of health insurance coverage frequently occurring with unemployment, the access problems of the uninsured are a pressing source of policy concern.

Simultaneously with the cutbacks in Medicaid, major reductions were made in the community and migrant health center programs. Overall funding was reduced by 25% in absolute dollars, which may lead to 1.1 million fewer people being served than the 6 million served in 1980. The National Health Service Corps, while not as seriously affected now, will be substantially reduced in future years, since no new scholarships are being awarded with commitments for service in underserved areas (Davis, 1981).

Financial strains on public hospitals and clinics supported by state and local governments are leading to further curtailment of services. Preadmission deposits, often sizeable amounts, impose serious barriers for many of the uninsured seeking hospital care. Teaching hospitals that have for years maintained an open-door policy are reevaluating the fiscal viability of continuing such a policy. In many areas, hospitals are beginning to transfer nonpaying patients to public facilities, further expanding the charity load of those facilities and reducing their ability to remain solvent (Brazda, 1982).

Recent hearings have documented the refusal of community hospitals to take uninsured patients, even in emergency situations. This has led to documented cases of deaths that could have been avoided with prompt medical attention (Committee on Energy and Commerce, 1981).

Recommendations

Such disparities in access to care are unacceptable in a decent and humane society. Several actions are required to assure progress toward adequate access for all. Medicaid coverage should be expanded to provide basic insurance coverage for all low-income individuals. Continued funding and expansion of the community and migrant health center programs to assure physical access to services for residents of high-poverty, medically underserved communities is an essential adjunct to broadened financing for low-income populations. As the metropolitan and nonmetropolitan differentials among the insured demonstrate, financing alone is not enough to correct access differentials. Resource development must be coupled with improved financing in underserved areas to assure that the needed providers are available. Other important ways to provide expanded insurance coverage without perpetuating the cost inefficiencies of the existing system include reform of Medicaid, Medicare, and private health insurance plans to encourage ambulatory care in cost-effective primary care programs, and experimentation with Medicaid capitation payments to primary care centers or networks of primary care centers.

References

Brazda, J. (Editor). 1982. *Perspectives: Who Will Care for the Uninsured?* Washington Report on Medicine and Health (September 27). Washington, D.C.: McGraw-Hill.

Committee on Energy and Commerce, U.S. House of Representatives. 1981. Hearings on Medicaid Cutbacks on Infant Care (July 27). Washington, D.C.: U.S. Government Printing Office.

Davis, K. 1982. *Medicare Reconsidered.* Presented at Duke University Medical Center, Private Sector Conference on the Financial Support of Health Care of the Elderly and the Indigent (March 14-16).

Davis, K. 1981. *Reagan Administration Health Policy.* Journal of Public Health Policy (December).

Davis, K. 1975. National Health Insurance: Benefits, Costs and Consequences. Washington, D.C.: The Brookings Institution.

Davis, K. and Schoen, C. 1978. Health and the War on Poverty: A Ten Year Appraisal. Washington, D.C.: The Brookings Institution.

David, K., Gold, M., and Makuc, D. 1981. *Access to Health Care for the Poor: Does the Gap Remain?* Annual Review of Public Health. Annual Reviews, Inc.

Department of Health and Human Services. 1982a. National Medical Care Expenditure Survey, 1977 (unpublished statistics). Hyattsville, Md.: National Center for Health Services Research, U.S. Department of Health and Human Services.

Department of Health and Human Services. 1982b. HEALTH CARE FINANCING TRENDS (June). Baltimore, Md.: Health Care Financing Administration, U.S. Department of Health and Human Services.

Grossman, M. and Goldman, F. 1981. *The Responsiveness and Impacts of Public Health Policy: The Case of Community Health Centers*. Presented at 109th Annual Meeting of American Public Health Association, Los Angeles, California (November).

Hadley, J. 1982. MORE MEDICAL CARE, BETTER HEALTH? Washington, D.C.: The Urban Institute.

Kasper, J.A., Walden, D.C., and Wilensky, G.R. 1978. *Who are the Uninsured? National Medical Care Expenditures Survey Data Preview Number 1*. Hyattsville, Md.: National Center for Health Services Research, U.S. Department of Health and Human Services.

Rowland, D. and Gaus, C. 1981. *Medicaid Eligibility and Benefits: Current Policies and Alternatives*. Presented at August 1981 Commonwealth Fund Forum, Health Care for the Poor: What Can States do in the Eighties? Washington, D.C.: Center for Health Policy Studies, Georgetown University.

Taylor, A.K. and Lawson, W.R. 1981. *Employer and Employee Expenditures for Private Health Insurance. National Medical Care Expenditures Survey Data Preview Number 7*. Hyattsville, Md.: National Center for Health Services Research, U.S. Department of Health and Human Services.

Wilensky, G.R. and Walden, D.C., 1981. *Minorities, Poverty, and the Uninsured*. Presented at the 109th Meeting of the American Public Health Association, Los Angeles, California (November). Hyattsville, Md.: National Center for Health Services Research, U.S. Department of Health and Human Services.

The Medicaid Program in Transition

Gerald R. Connor, M.C.R.P.*

Access by the poor to basic services has always been a major philosophical and practical problem for democratic societies. A recent analysis of this long-standing debate as it concerns medical care begins by defining a principle of equal access to health care that:

> . . .demands that every person who shares the same type and degree of health need must be given an equally effective chance of receiving appropriate treatment of equal quality so long as that treatment is available to anyone.[1]

The author concedes that this ideal creates an unpleasant choice for government in a democracy: Either the government must provide the most expensive care to all who cannot otherwise afford it or it must restrict the access of the most well-off to some of the most expensive care. Neither choice is likely in the United States. As a practical matter, she concludes that an expanding supply of health services and a significant government commitment to provide health care services to all will both satisfy democratic principles and move society toward the equal-access ideal.

Medicaid and Medicare, among other programs, can be viewed as attempts at the margin to approach the principle of equal access in our society.[2] Through these programs, two major disadvantaged segments of our society were guaranteed access

*Health Policy Consultant, Washington, D.C.

April 1982

[1] Amy Gutmann, *For and Against Equal Access to Health Care*, Milbank Memorial Fund Quarterly 59:4 (Fall 1981), p. 542.

[2] *Ibid.*, p. 557. Gutmann argues that Medicaid is flawed as a democratic solution because it does not bind all citizens. As a result it is, she suggests, a more likely candidate for withdrawal of political (and therefore budgetary) support.

to at least basic medical care. Both programs—and other public policy initiatives—also seek to insure the elasticity of the supply side of the equation by paying the "going rate" for services or by directly increasing the supply of services (e.g., through the provision of financial assistance for health professions education.)

This paper will describe the Medicaid program, designed to increase access by the poor to medical care. First, the public policy environment in which the program was born is described. Next the structure of the program is described, followed by a summary of the differences among state Medicaid programs prior to the enactment of amendments to the Federal law in 1981. The next section of the paper describes the changes to Federal Medicaid law enacted as part of the Omnibus Budget Reconciliation Act. The fifth section of the paper discusses changes to their Medicaid programs being contemplated by state governments at the beginning of fiscal year 1982 (FY 1982). Finally, the paper outlines several issues that can affect access to medical care and that may appear on the agendas of Medicaid decisionmakers during 1983.

Background

Through most of the history of the United States, the delivery of medical care to the poor has been viewed as a local matter. Private charity and the local public hospital were the traditional responses, with state government assuming responsibility for institutional care for mental illness and chronic physical illness. During the Depression, the Federal government first made grants available to state governments for care for unemployed persons, through the Federal Emergency Relief Administration.

The Social Security Act, passed in 1935, provided cash assistance to certain categories of needy persons (the aged, blind, disabled, and families with dependent children). The assistance provided by state governments (with Federal financial assistance) was assumed to cover the costs of food, shelter, clothing, and medical care for the needy poor. By 1950, it was clear that the cash assistance authorized by the Social Security Act was inadequate to cover basic living costs and unplanned medical expenses.

In 1950, the law was amended to allow Federal financial participation in state medical "vendor payment" programs. States could purchase medical care services on behalf of cash assistance recipients directly from providers, and the Federal government would share in the costs. By 1960, 40 states had established vendor payment programs and total expenditures had reached approximately $500 million per year.

In the long debate over national health insurance in the United States, the needs of the elderly had always provided the most compelling argument for some form of public intervention. The specter of declining income and health for a growing elderly segment of the population was especially difficult to ignore in the post-World War II years. In 1960, a bill sponsored by Senator Kerr and Representative Mills proposed to expand state vendor payment programs to cover aged persons who were not poor enough to qualify for cash assistance but were too poor to afford adequate medical care. The Medical Assistance for the Aged (MAA) program became law, and by 1965 total expenditures under state vendor payment programs grew to $1.3 billion.

The Federal share in state vendor payment programs had increased from a limit of $15 per month per aged recipient to encompass 50-80% of the states' expenditures under MAA. By 1962, however, only 28 states had begun MAA programs, and the Senate Special Committee on Aging had concluded "there is no reason to expect that it will become a [national program] in the foreseeable future." The foremost policy concern in the early sixties, then, was to find a more effective means to guarantee medical care for the needy aged.

In 1965, the mood of Congress and the considerable political skills of President Lyndon Johnson combined to enable a major revision of the Social Security Act's provisions for medical care for the elderly. The 1965 amendments (by adding Title 18 to the Act) created the Medicare program, which financed hospital care (Part A) through a payroll tax and ambulatory care (Part B) through a voluntary premium paid by the beneficiary. Almost all Americans aged 65 and over would be eligible for Medicare coverage.

The 1965 amendments also adjusted the payroll tax to finance retirement and disability benefits and added Title 19 to the Social Security Act, consolidating the state medical care vendor payment programs authorized under various other titles. Title 19, according to the Senate Finance Committee, was designed "to provide a more effective Kerr-Mills program for the aged and extend its provisions to additional needy persons." The Senate bill, in fact, would have allowed the states to choose to continue their existing programs or to establish a new Medicaid program for all their welfare-eligibles.

It was estimated that the first year's operation of the new Title 19 program would cost an additional $238 million. In 1966, vendor payments were exactly $238 million more than in 1965—but only six states had initiated Medicaid programs. It proved to be the only year for which accurate estimates could be made: For FY 1968, with 37 states participating, the cost estimates included in the President's budget ($2.25 billion) were 35% below the actual expenditures for the year ($3.45 billion).

From a rather modest beginning as a consolidation of state vendor payment programs, Medicaid has become the largest single item in state government human services budgets and a major Federal welfare program. Table 1 shows the growth in Medicaid expenditures from its beginning through 1982. Both levels of government are constantly seeking rational and humane ways to contain its growth.

TABLE 1: Total Medicaid Vendor Payments (State and Federal)

Fiscal Year	Expenditures (billions)	Change (%)
1966 (baseline)	$ 1.6	0
1967	2.3	+42.7
1968	3.5	+52.0
1969	4.4	+26.1
1970	5.1	+17.1
1971	6.3	+24.6
1972	7.3	+15.8
1973	8.7	+18.6
1974	9.7	+11.7
1975	12.1	+24.1
1976	14.0	+15.6
1977	16.4	+17.0
1978	18.0	+ 9.8
1979	20.5	+13.7
1980	23.3	+13.9
1981*	28.5	+22.3
1982*	31.7	+11.3

*Estimates
Source: *Data on the Medicaid Program—Eligibility/Services/Expenditures*, published from time to time by the U.S. Department of Health and Human Services. Figures are from the 1979 edition (revised) and have been updated through conversations with Department staff preparing the 1981 edition.

The Medicaid Program

The Medicare program enacted into law a form of social insurance under which the working population paid a premium (in the form of a payroll tax) in return for benefits when they were no longer working. (The elderly paid a premium directly for ambulatory care benefits under Part B of the program.) It was, and remains, a national program with clear eligibility rules and covered benefits. It is administered by the Federal government uniformly across the nation.

Medicaid, on the other hand, consolidated and expanded a variety of state-administered vendor payment programs. Eligi-

bility (within broad Federal guidelines) is established by states under varying standards; the benefit package (beyond the basic services required under the law) is designed and from time to time adjusted by state government. In effect, Title 19 created 53 separate programs not based on insurance principles and dependent to a large extent upon the availability of public revenues.

Measurement of these programs is an inexact science. Within the confines of Title 19 of the Social Security Act, there are 54 major decisionmaking units (53 state and territorial programs plus the Federal government). Each major unit has a number of constituent parts. Data are collected, arrayed, and analyzed to suit the needs of each decisionmaker. During the fall of 1981, for example, there were three "official" estimates of the Federal cost of the Medicaid program for fiscal year 1982: One was prepared by the Federal Office of Management and Budget, one by the U.S. Department of Health and Human Services, and the third by the Appropriations Committees of the U.S. Congress. At one time, there was a billion-dollar difference among the estimates; at no time did they agree.

The following characteristics of the Medicaid program determine the type of data collected:

(1) The program is designed as a vendor payment mechanism, and the basic unit of analysis is the payment of a provider for a service.

(2) The transaction between a state program and a provider presumes that the recipient of the service is eligible for Medicaid and that the service is covered. Both presumptions can be challenged by the Federal government.

(3) The very size of the program (at least 20 million persons receive some services each year) means that even a small percentage of challenged transactions generates an extremely large number of sometimes protracted negotiations between Federal and state governments.

These basic characteristics of the program mean that data on people served will focus on *recipients* rather than on *eligibles*. Eligibility only becomes a Medicaid question when services are used; people who are eligible but do not use Medicaid are not counted by a system designed to pay for services. The resolution of Federal challenges to state reimbursements under Medicaid can be time-consuming. The resolution does not necessarily come during the fiscal year in which the service was rendered—nor, in some cases, during the succeeding fiscal year. "Final" Medicaid expenditure data are in fact revised each time an agreement is reached. The nature of the program thus also insures that final data will not be available until long after the close of a given fiscal year. In this paper, an effort has been made to present the most current data available.

State Fiscal Capacity. Most discussions of national health policy focus on the taxing and spending capacity of the Federal government to bear the policy burden. The responsiveness of Federal revenues to major economic trends is dampened somewhat by the size of the revenue base: A major recession in the automobile industry in Michigan can be offset to some degree by the booming energy business of Texas and Colorado. Even when the national economy does not perform up to expectation, the Federal government can borrow on future revenues to maintain a steady course for national spending policy. In most years we can expect national policy, financed with national revenues, to remain fairly consistent.

As Table 2 demonstrates, state government tax revenues move in parallel with one major indicator of the health of the national economy. As the gross national product (GNP) rises rapidly, state revenues do as well; when the GNP decelerates, the impact on state tax revenue is direct and immediate. Federal tax revenues, on the other hand, can be a tool of national economic policy as well as a result of it: When the GNP declines, Federal revenues can decelerate even more rapidly, allowing higher consumption in the private sector designed to prime the sagging national economy; when the national economy heats up, Federal tax collections tend to rise even more rapidly.

TABLE 2: Percentage Increase in Tax Revenues and Gross National Product

Year	Gross National Product (GNP)	Federal Tax Revenue	State Tax Revenue
1972	9.2	2.6	11.7
1973	11.6	7.4	13.7
1974	8.1	11.6	9.0
1975	8.6	3.2	8.0
1976	11.3	6.0	11.4
1977	11.6	21.1	13.3
1978	12.0	12.6	12.0

Source: U.S. Department of Commerce, STATISTICAL ABSTRACT OF THE UNITED STATES: 1980 (Washington: Government Printing Office).

During the time period covered by the data in Table 2, there were no major changes in Federal policy governing state Medicaid programs. During that same period, however, state governments were forced to react quickly to changing economic conditions in order to control state budgets, which by state law or constitution must be in balance. To cope with rising costs in

their largest program, states can adjust Medicaid eligiblity, the benefits promised by the program, and the level of reimbursement for services provided under the program.

Eligibility. The Federal law goverening Medicaid requires that eligibility be granted to everyone within a state who receives cash assistance under one of the titles of the Social Security Act. These "categorically needy" persons include the aged, blind, disabled, and members of families with dependent children who meet State low-income standards. Eligibility for The Aid to Families with Dependent Children (AFDC) program—and thus for Medicaid—is determined by state government. Cash assistance programs for the aged, blind, and disabled were combined in 1972 under the federally-funded Supplementary Security Income (SSI) program. Medicaid eligibility for recipients of SSI cash assistance is determined by the Federal government for 29 states.

Each state's Medicaid plan must make all AFDC and SSI recipients eligible.[3] Each state may, at its option, grant eligibility to other needy persons whose income or family composition is such that they are not eligible for cash assistance. These "categorically-related needy" or "medically needy" persons are generally those who would be covered under the broadest state public asistance program allowed under Federal law and whose income does not exceed 133.33% of the state's present welfare standard.

Presently, every state except Arizona[4] has implemented the basic Medicaid program. By 1979, 30 state programs (including the District of Columbia) had expanded to cover the medically needy. Eligibility standards established by the states resulted in more than 12 million recipients[5] of Medicaid-reimbursed services in 1969. By 1975, the recipient count had exceeded 21 million persons, peaking at 23.5 million in 1976, and recently stabilizing at approximately 22 million persons.

---

[3] Approximately 20 states take advantage of a provision of the 1972 law creating SSI to limit Medicaid eligibility. States are permitted, under that provision, to limit SSI eligibility for Medicaid to those persons who would satisfy the state's eligibility standards for cash assistance to the aged, blind and disabled in effect prior to the enactment of the Federal program.

[4] Arizona does not have a Medicaid program but has begun negotiations seeking a three-year experiment with Federal financial support. The state proposes to contract for medical care for the poor on the basis of a prepaid capitation with a variety of insurers and providers. The experiment would provide all Medicaid services except those of skilled nursing facilities and would entail a Federal share of 55%.

[5] No accurate count of Medicaid-eligibles exists, because eligibility is certified upon application on a case-by-case basis. A count of recipients is the most commonly used surrogate.

Generally, Medicaid eligibility increased significantly during the 1970s. To some extent, this increase mirrors an increase in welfare eligibility; but the variation in recipient counts for the two programs is not directly proportional from year to year. (See Table 3). To some extent, the increase reflects expansion of eligibility in state programs. (Between 1970 and 1975, for example, six states added coverage of the medically needy to their programs.)

TABLE 3: Percentage Changes* in Medicaid and Cash Assistance Recipients

Fiscal Year	Medicaid Recipients	Cash Assistance Recipients
1972	+ .1	+9.2
1973	+ 4.6	+3.1
1974	+10.7	+ .1
1975	+ 1.7	+6.0
1976	+10.7	+3.6
1977	− 2.8	−1.4
1978	− 4.5	− 1.0
1979	− 1.2	+ 3.4
1980	+ .3	+ 2.5
1981	+ 4.0	NA

* From previous year.

Source: *Data on the Medicaid Program—Eligibility/Services/Expenditures*, published from time to time by the U.S. Department of Health and Human Services. Figures are from 1979 edition (revised) and have been updated through conversations with Department staff preparing the 1981 edition.

States, like the Federal government, find it relatively easy to expand eligibility for Medicaid. Most attempts to restrict eligibility levels, on the other hand, are met with public disagreement and strong political resistance.

Benefits. In order to qualify for Federal matching payments under Federal law, a state Medicaid program must provide at least:

(1) Hospital services (outpatient and inpatient);

(2) Physician's services;

(3) Diagnostic services (laboratory and x-ray);

(4) Family planning services;

(5) Nursing home services (in skilled nursing facilities);

(6) Children's screening diagnosis and treatment services;

(7) Rural health clinic services; and

(8) Nurse midwife services.

These basic services must be provided to all categorically needy persons in the state (i.e., those receiving cash assistance under AFDC and SSI). States may, at their option, include reimbursement for any other medical care service recognized by state law. Among the optional services, every state covers the services of intermediate care facilities and all but two states cover prescription drugs.

If a state granted eligibility to the medically needy prior to FY 1982, it was required to provide either the basic services or any seven Medicaid services. All states with a medically needy program offer the basic services and a wide range of optional services to this population.

The structure of the Medicaid benefit package is weighted toward institutional care. This weighting is clear from an examination of the proportion of Medicaid expenditures going to facilities. (See Table 4.) Since 1972, Medicaid expenditures for hospitals and long-term care facilities (skilled nursing and intermediate care institutions) consumed about 70% of the budget, while the proportion spent on physician services has declined steadily. Long-term care is the largest and fastest growing component of Medicaid costs: In FY 1977, Medicaid reimbursement accounted for 50.8% of nursing home revenues nationwide ($6.4 billion).

The Federal Medicaid statue allows states to impose reasonable limits on the amount, duration, and scope of the serv-

TABLE 4: Percentage Distribution of Medicaid Expenditures

Year	Hospital Services	Long-Term Care Facilities	Physician Services
1972	37.9	31.5	11.3
1973	34.8	35.0	10.7
1974	33.0	36.0	10.8
1975	31.0	37.9	10.2
1976	31.6	37.3	9.8
1977	31.5	39.2	9.2
1978	30.8	41.9	8.8

Source: *Data on the Medicaid Program—Eligibility/Services/Expenditures*, published from time to time by the U.S. Department of Health and Human Services. Figures are from 1979 edition (revised) and have been updated through conversations with Department staff preparing the 1981 edition.

ices (including the basic services) for which coverage is provided. In 1979, for example, 18 states limited the number of inpatient hospital days for which reimbursement was available. The limits varied from 20 days per year in Alabama and Tennessee to 60 days per spell of illness in Ohio. In addition, 14 states require prior authorization by the state Medicaid agency for some kinds of acute hospital care.

Reductions in the benefit package and limits on the amount, duration, and scope of those services are less difficult to accomplish than reductions in eligibility. In spite of that relative ease, between 1970 and 1975 almost every state increased the number of optional services available through its program, and the number of states with no limits on basic services increased from 9 to 30.[6]

Reimbursement and Costs. Federal Medicaid law requires states to establish a method for reimbursing providers that is approved by the Secretary of Health and Human Services. The reimbursement method for hospitals was tied, until this year, to payment on the basis of reasonable costs as determined by the Federal government under Medicare. Most other providers (e.g., physicians, laboratories) are paid on the basis of a fee—usually the prevailing customary fee charged by similar providers in the area.

The Medicaid and Medicare statutes were quite clear about the need to pay reasonable costs or charges. It was feared that the promise of access to "mainstream" medical care for the poor and elderly would be meaningless if the programs could not convince most providers to treat their beneficiaries. If reimbursement rates were set too low, policymakers were convinced that the new programs would simply perpetuate the two-class system of care that existed prior to 1965.

From the beginning, Federal approval of state reimbursement practices has been a major focus of debate. National associations of providers have, from time to time, taken exception to state definitions of "reasonable" costs and charges. Federal administrators have had to balance program costs, state fiscal capacity, and access by the poor to medical care in reviewing state Medical Assistance Plans and proposed amendments.

Through the years, a balance apparently has been maintained. In 1964, almost 30% of low-income persons in the United States had not seen a physician during the previous year, compared with less than 18% of those above the poverty line.[7] By 1977, less than 15% of low-income persons and 13% of

[6] Stephen M. Davidson, *Variations in State Medicaid Programs*, JOURNAL OF HEALTH POLITICS, POLICY AND LAW 3:1 (Spring 1978), pp. 54-70.
[7] National low-income and poverty standards used in this paper refer to the standard developed in 1964 by the Social Security Administration based on an economy budget for food. Since 1969 the standard

those more well-off had not visited a physician during the previous year.

By the early 1970s, Medicaid costs had begun to rise at a rate two or three times the increase in consumer prices generally. (See Table 5.) This increase could not be attributed to parallel increases in Medicaid recipients, and another reimbursement issue came to the fore. National expenditures for hospital care had begun to increase at rates well above the inflation of costs in most other sectors of the economy. Reimbursement of the "reasonable costs" of providing hospital care by Medicare, Medicaid, and some private insurers had become part of this new problem. If a hospital was assured of recovering its costs, it had no incentive for economy; policymakers began to search for ways to change the reimbursement system.

TABLE 5: Percentage Changes in Consumer Prices and Medicaid Expenditures

Year	Consumer Price Index (CPI)	Medicaid Expenditures	Medicaid Recipients
1972	3.3	15.8	.1
1973	6.2	18.6	4.6
1974	11.0	11.7	10.7
1975	9.1	24.1	1.7
1976	5.8	15.6	(−2.8
1977	6.5	17.0	(−4.5)
1978	7.7	9.8	(−1.2)
1979	11.3	13.7	.3)
1980	13.5	13.9	4.0
1981	10.4	22.3	

Source: *Data on the Medicaid Program—Eligibility/Services/Expenditures*, published from time to time by the U.S. Department of Health and Human Services. Figures are from 1979 edition (revised) and have been updated through conversations with Department staff preparing the 1981 edition.

Several state governments began programs to review and approve hospital budgets and costs at the beginning of the year. Hospital reimbursement was based upon these prospective budgets, and excessive reimbursement was denied. The state programs varied from Rhode Island's negotiation of the allowable increase in hospital budgets to the establishment of New Jersey's system, which based reimbursement for each hospital stay on the diagnosis of the patient upon admission. As many

has been raised annually to account for changes in the Consumer Price Index.

as nine states had established some form of mandatory rate review by 1980.[8]

Most state efforts at hospital cost containment regulated rates in a manner not envisioned by the law and regulations, which were intended to guarantee "reasonable cost" recovery by hospitals. As a result, states were forced to seek exemptions from these provisions in order to apply their rate-setting mechanisms to Medicare and Medicaid reimbursement. In the absence of an exemption, state rate regulation could only apply to the bills sent to private insurers and to patients who paid their own bills.

In 1977, the Carter administration began a four-year effort to enact a national hospital cost-containment statute. The proposal would have exempted hospitals in states with pre-existing rate review programs and would have imposed a national ceiling on allowable hospital cost increases. Two successive Congresses were unable to agree on a hospital cost-containment bill.

State Medicaid Programs

In August 1981, every state (except Arizona) and four territories had chosen to accept the offer of Federal financial assistance for Medicaid. In that month, the most significant changes in the Federal standards governing the program became law. The Medicaid programs in effect prior to the 1981 amendments reflected vast differences among the states in fiscal capacity and political philosophy. (See Table 6.)

Thirty states have chosen to grant eligibility to persons with incomes too high to qualify them for cash assistance but too low to cover their medical expenses (the medically needy). The decisions by state governments to expand or contract eligibility for their Medicaid programs have, during the past decade, tended to increase the number of recipients.

Medicaid eligibility is not simply a function of need measured by income. Eligibility standards established in Federal law and modified by state decisions result in significant variations. Most Medicaid-eligibles fall into one of the Federal categories. (One- and two-adult households, for example, are not usually covered by state programs since they do not fit a categorical definition.) Since states set the income standards for eligibility, persons below the national poverty standard in some states are not eligible, and persons in other states above the national standard are eligible. In states offering coverage to the medically needy, extraordinary medical expenses can be deducted from income in order to determine Medicaid eligibility (poten-

[8] See: Michael Zubkoff, Ira E. Raskin, Ruth S. Hanft, eds., HOSPITAL COST CONTAINMENT (New York: Milbank Memorial Fund, 1978).

tial recipients can "spend down" to the medically needy income level). As a result, only 35% of the persons living in poverty in 1980 were recipients of Medicaid services. Stated another way, 48% of the people who were served by Medicaid in 1980 (10.4 million persons) were not living below the national poverty standard.[9]

It is difficult to make accurate predictions of Medicaid eligibility based on past trends in the number of Medicaid recipients. Some observers have concluded, on the basis of those trends, that Medicaid program growth peaked in 1977 and is declining. More recent data on Medicaid recipients indicates that eligibility continues to expand, albeit at a slower rate.[10]

States must also make choices among the nonmandatory services to be covered by their programs. By the end of the last decade, 34 states chose to offer at least 10 optional services to their categorically eligible Medicaid population. (Of the states covering medically needy persons, 22 offer at least 10 of the 17 optional services to this optional group of eligibles.)

While there are differences in the reported expenditures by states for the services provided by their Medicaid programs, the differences are not solely the result of reimbursement decisions. Expenditures can vary because of eligibility decisions (adding eligible children will cost less than adding eligible elderly persons), and because of service decisions (restricting the number of covered days of acute hospital care will reduce expenditures significantly). An analysis of the amount of the difference attributable to reimbursement decisions alone is beyond the scope of this paper.

It is clear that the tendency has been to expand the scope and coverage of state Medicaid programs between 1965 and the enactment of the 1981 amendments. That tendency, however, cannot mask the vast differences among the programs:

(1) Twenty states do not offer Medicaid to persons ineligible for cash assistance but unable to meet large medical expenses.

---

[9] Data on Medicaid coverage of persons below the national poverty standard for 1980 were tabulated in: U.S. Department of Health and Human Services, NATIONAL HEALTH INSURANCE WORKING PAPERS, Vol. 2: DECISION PAPERS FOR THE SECRETARY, undated (Washington: U.S. Department of Health and Human Services).

[10] See, for example, the testimony of Clifton Gaus before the Commission. In 1978, recipient counts began to decline for the first time, but more recent data indicate that the steady, gradual increase has resumed. The administration predicts a minor decrease in recipients in fiscal year 1983 if its proposals (including further reductions in AFDC eligibility) are enacted by the Congress.

TABLE 6: Comparison of State Medicaid Programs and Planned Changes

State	Medically Needy Program	Number of Optional Services for Categorically Eligible	Number of Optional Services for Medically Needy
Alabama	No	9	0
Alaska	No	8	0
Arizona (no program)	0	0	0
Arkansas	Yes	13	11
California	Yes	16	16
Colorado	No	9	0
Connecticut	Yes	15	15
Delaware	No	7	0
Dist. of Columbia	Yes	14	14
Florida	No	6	0
Georgia	No	8	0
Hawaii	Yes	11	11
Idaho	No	8	0
Illinois	Yes	17	17
Indiana	No	16	0
Iowa	No	12	0
Kansas	Yes	16	16
Kentucky	Yes	11	11
Louisiana	Yes	10	5
Maine	Yes	12	3
Maryland	Yes	10	9
Massachusetts	Yes	15	15
Michigan	Yes	14	14
Minnesota	Yes	17	17
Mississippi	No	6	0
Missouri	No	7	0
Montana	Yes	15	15
Nebraska	Yes	15	15
Nevada	No	13	0
New Hampshire	Yes	14	13
New Jersey	No	16	0
New Mexico	No	12	0

Percentage of Low-Income Population Receiving Medicaid	*Changes Adopted or Considered:		
	Eligibility	Services	Reimbursement
34	–	–	–
19	0	0	0
0	0	0	0
28	0	–	–
39	–	+/–	–
30	+	+	+
36	–	–	–
34	0	0	–
53	NR	NR	NR
27	+	+	0
36	–	–	–
30	–	–	–
23	0	+	–
43	–	–	–
26	+	–	0
25	+	–	–
29	–	–	–
27	–	–	–
35	–	–	–
30	0	0	–
45	–	–	–
46	0	+	–
42	–	–	–
26	–	–	–
36	+	–	+/–
32	–	–	–
21	–	–	–
20	–	0	+/–
20	0	–	–
33	0	–	–
38	0	–	–
27	–	–	–

TABLE 6: Comparison of State Medicaid Programs and Planned Changes (*Continued*)

State	Medically Needy Program	Number of Optional Services for Categorically Eligible	Number of Optional Services for Medically Needy
New York	Yes	14	14
North Carolina	Yes	12	12
North Dakota	Yes	15	15
Ohio	No	15	0
Oklahoma	Yes	7	7
Oregon	No	15	0
Pennsylvania	Yes	12	8
Rhode Island	Yes	9	8
South Carolina	No	11	0
South Dakota	No	9	0
Tennessee	Yes	8	8
Texas	No	9	0
Utah	Yes	15	15
Vermont	Yes	9	9
Virginia	Yes	11	11
Washington	Yes	17	17
West Virginia	Yes	13	13
Wisconsin	Yes	17	17
Wyoming	No	3	0
U.S. Summary:	30 Yes 20 No		

Sources: The Intergovernmental Health Policy Project, RECENT AND PROPOSED CHANGES IN STATE MEDICAID PROGRAMS: A FIFTY STATE SURVEY (Washington: May and October, 1981); *Data on the Medicaid Program—Eligibility/Services/Expenditures*, published from time to time by the U.S. Department of Health and Human Services (figures are from the 1979 edition, revised, and have been updated through conversations with Department staff preparing the 1981 edition); and U.S. Department of Health and Human Services, NATIONAL HEALTH INSURANCE WORKING PAPERS, VOL. 2: DECISION PAPERS FOR THE SECRETARY (Washington: DHHS, undated).

Percentage of Low-Income Population Receiving Medicaid	*Changes Adopted or Considered:		
	Eligibility	Services	Reimbursement
43	+/−	+/−	+/−
31	−	−	−
18	+	+	0
40	+/−	0	−
28	0	0	0
25	+/−	−	−
38	−	−	+
41	0	0	−
29	0	−	−
14	0	+	−
35	−	−	−
23	0	+	−
25	−	−	−
29	0	−	−
29	0	−	−
33	−	−	−
37	0	−	0
26	0	−	−
18	0	0	0
35	22 − 6 + 3 +/−	32 − 7 + 2 +/−	37 − 2 + 3 +/−

(2) Sixty-five percent of the people living in poverty in 1980 were not receiving Medicaid services.

(3) Sixteen states provide less than ten of the optional services to all Medicaid eligibles (for example, 19 states do not provide dental services and 21 provide no physical therapy services).

Before the 1981 amendments to the Federal law, Medicaid eligibility and coverage was, to a large extent, an accident of one's state of residence. The amendment (described in the following section) will require the states to reconsider the scope of their programs.

1981 Medicaid Amendments

As part of his 1981 economic recovery program, President Reagan proposed to reduce Medicaid expenditures by limiting the Federal share in FY 1982 to 5% more than it had been in FY 1981. In order to make the proposal more palatable to state governments, he proposed to relax many of the federally imposed standards for approval of state Medical Assistance Plans.

The Congress, instead, chose to reduce the Federal share otherwise payable to states by 3% in FY 1982, 4% in FY 1983, and 4.5% in FY 1984. A state can restore one percentage point of the reduction if it operates a qualified hospital cost-containment program, another point if it suffers an unemployment rate exceeding 150% of the national average, and a third percentage point if it recovers an amount equal to 1% of the Federal share through aggressive efforts to root out fraud and abuse. To help states make up for the loss in Federal funds, Congress agreed to relax many of the Federal Medicaid standards.

In order to restrain the increase in Medicaid expenditures, states were allowed (in some cases, forced) to make changes in eligibility, services, and reimbursement practices. As state legislatures meet, restructuring state Medical Assistance Plans will be high on their agendas.

Changes in Eligibility. States will still be required to grant Medicaid eligibility to all SSI and AFDC cash assistance recipients. The rules for determining AFDC eligibility, however, have been revised significantly. States must use the new eligibility criteria in order to qualify for Federal matching payments for AFDC and Medicaid expenditures.

Under Federal law governing AFDC eligiblity prior to FY 1982, states were allowed to consider a variety of family circumstances in determining eligibility. States could exclude the assets of each recipient up to $2000 and were permitted to exclude any income that could reasonably be considered as work expense (e.g., child care expenses). There was no limit on the amount of income that could be disregarded in determining

eligiblity. States could limit eligibility to children under 18 years of age or could cover all children who were students and under the age of 21. States could not count the income of step-parents in determining a child's eligibility unless, by state law, the income was available to the child. Eligibility could be granted to intact families if either parent was unemployed.

Each of these permitted practices was changed by amendments enacted as part of the Omnibus Budget Reconciliation Act of 1981. Beginning on October 1, 1981, states are prohibited from exempting more than $1000 in assets per family and may not consider any family earning more than 150% of the state standard of need (before disregards) for AFDC eligibility. The amount of income that can be disregarded as a work expense is limited to $75 per month, in addition to a maximum of $160 per child for day-care services. The income of step-parents must now be counted as available to the child in determining eligibility. States may no longer grant AFDC eligibility to an intact family unless the parent with the higher income is unemployed. The changes in AFDC eligibility standards are summarized in Table 7.

TABLE 7: Changes in AFDC Eligibility Standards

Standard	Before October 1, 1981	After October 1, 1981
Exempt resources	$2000 per person	$1000 per family
Maximum gross income	No limit	150% of state standard
Income disregards	All reasonable work expenses	$75 maximum + $160 per child day care
Students	All under age 18, or all under 21	All under age 18 or 19 or 20 or 21
Step-parent income	Counted only if available to child	Counted
Unemployed parent	Either parent unemployed	Higher-earning parent unemployed
Medically needy (optional)	All categories covered	Some categories covered

As an example of the impact of these changes on the complex structure of Federal and state eligibility standards, the Children's Defense Fund points out:

> . . .in Texas the standard of need for a family of four is $187 per month. Currently, AFDC earned income disregards and work expense deductions designed to encourage employment can leave some families with a total income, including earnings and partial AFDC grant, in excess of 150 percent of the State's standard of need . . . Once the 150 percent cap takes effect, that same family will be completely ineligible—despite the fact that their income is only 40 percent of the Federal poverty guidelines.[11]

Admittedly, Texas's standard of need is exceptionally low; but in 1978, 11 other states had standards below $300 per month for a family of four.

States are now permitted to grant AFDC eligibility to students up to age 21 or to any age group between 18 and 21 years of age. In the past, states that refused AFDC eligibility to children between 18 and 21 because they did not meet age or student requirements were nevertheless required to grant Medicaid eligibility to those children. An amendment included in the Reconciliation Act removes this requirement.

Prior to fiscal year 1982, states choosing to implement a medically needy program were required to include all categorically related needy persons (aged, blind, disabled and families with dependent children). States may now provide eligibility to some portion of the medically needy population without extending coverage to any other category of individuals.

Most observers feel that the eligibility changes can have the most significant impact on Medicaid costs and coverage. Potentially, a large number of low-income people will no longer be eligible for AFDC cash assistance and will therefore lose their Medicaid eligibility.

Changes in Benefits. States must continue to provide the basic services to all categorically needy people. The major changes in service coverage enacted with the Omnibus Budget Reconciliation Act concern (1) the state's ability to restrict the patient's freedom to choose a provider and (2) the services that must be provided to anyone eligible under the state's medically needy program.

Prior to FY 1982, all Medicaid recipients were guaranteed the right to seek medical care services from any qualified provider within the state. States felt this right was meaningless in some cases (e.g., laboratory services) and abused by some recipients who "shopped around" (i.e., sought treatment from several providers for the same illness). The 1981 amendments allow states to purchase laboratory services and medical devices for all Medicaid recipients from one or more suppliers under a competitive bidding arrangement. In addition, states may "lock in" a Medicaid recipient to a single provider if it can be shown that the recipient is overutilizing services without

[11] Children's Defense Fund, CDF Reports 3:7 (September 1981), p. 1.

cause. The Reconciliation Act allows the Secretary of Health and Human Services to waive—under a simplified waiver process—the freedom of choice requirement under certain other circumstances that benefit the program while not restricting access to Medicaid services.

In the past, states choosing to establish a medically needy program were required to provide at least seven of any Medicaid services to all medically needy eligibles throughout the state. Under the 1981 amendments, states need not cover all medically needy categories and may vary the services provided to each covered group. Services need not be comparable in amount, duration, and scope among categories of medically needy but must include ambulatory services for any group for whom institutional services are provided. In addition, states maintaining any form of a medically needy program must cover some additional children and pregnant women (i.e., categorically related children and categorically eligible women who fall within the medically needy income standard). The state must provide at least ambulatory care services for low-income children and prenatal and delivery services for pregnant women under this provision of the new law.

In general, the benefit changes permitted by the 1981 amendments are designed to give state administrators a means to control some of the costs associated with their Medicaid programs. It is likely that some of those remaining eligible for Medicaid will experience some limits on the services and types of providers to which they have access. The Federal law clearly frowns on limits that deny access but, in the final analysis, no conclusions can be drawn before the state legislatures exercise their new options.

Changes in Reimbursement. Since 1965, reimbursement of hospitals under state Medicaid programs had been tied to the "reasonable costs" defined by the Federal Medicare program. States were allowed to develop their own reimbursement methods (after 1967, the Secretary was encouraged by law to promote reimbursement experiments), but hospital reimbursement was always tied to Medicare costs.

Since the mid-1970s, several states have received long-sought approval to include Medicare and Medicaid reimbursement within their hospital rate and budget review schemes. The vast majority of hospitals, however, operate in states required to reimburse on the basis of Medicare costs.

The 1981 amendments to the Social Security Act repealed the requirement that Medicaid hospital reimbursement methods be designed to cover Medicare-defined costs. State reimbursement for hospital inpatient services now must be "reasonable and adequate to meet the costs which must be incurred by efficiently and economically operated facilities. . ."[12]

[12] See: Section 1902(a)(13)(A) of the Social Security Act.

Possible State Actions Under 1981 Amendments

In order to make the changes permitted by the 1981 Federal Medicaid amendments, in almost every case the state agency administering the program must seek changes in state law. The legislatures of 42 states will meet in 1982, with 74 of the sessions beginning on or before February 9.

In the recent past, most states have been expanding their Medicaid programs. A recent study constructed an index of state Medicaid programs that measures changes in eligibility, services, and reimbursement. Between 1970 and 1975, the index for only 3 states declined, while 37 states took actions that expanded the coverage of their programs.[13] Another analysis of state decisions that modified eligibility for Medicaid found 255 such actions between 1978 and July 1981. Of those decisions, 222 increased the number of persons eligible for the program.[14] While no estimates of new eligibles or new service dollars are available, these independent observations indicate a general expansionary trend for the programs during the past decade.

Now state governments will be forced to act to maintain Medicaid eligibility at nearly its present level. Three of the changes in Federal eligibility standards cannot be affected by state decisions: States may not exempt more than $1000 of resources per family, must count the income of step-parents, and may not use the unemployment of the lower-paid parent to determine AFDC eligibility for federally matched cash assistance. On the other hand, state programs could be modified to mitigate the remainder of the new Federal eligibility standards. Texas, to return to the earlier example, could raise its standard of need to a level that would maintain present eligibility based on the 150% maximum. States also have the option of retaining eligibility for those groups (e.g., students) whose coverage was previously mandatory.

When the United States Congress began to debate reduced Federal financial participation in Medicaid, states were forced to begin planning for this contingency. States that already had begun to think of retrenchment because of the downturn in the national economy now must also make orderly adjustments to accommodate a variety of changes in Federal policy. The Intergovernmental Health Policy Project[15] and the National Gover-

---

[13] Davidson, *loc. cit.*, p. 65.

[14] L. Bartlett, National Governors' Association, personal communication, January 1981.

[15] Intergovernmental Health Policy Project, Recent and Proposed Changes in State Medicaid Programs: A Fifty State Survey (Washington: May and October 1981). The National Governors' Association now maintains a computer file of state program changes for this ongoing project.

nor's Association have monitored the contingency plans throughout 1981.

Eligibility. In October 1981, 22 states were considering or had adopted proposals that would reduce Medicaid eligibility (see Table 6). The uncertainty surrounding Federal changes led to consideration in North Carolina (whose legislature does not meet in 1982) of a proposal to allow the Medicaid agency to deny eligibility to students and the medically needy, if necessary to meet the state's budget. Four states dropped coverage of intact families (with an unemployed parent) from their AFDC programs and 4 States adopted more restrictive eligibility standards for their medically needy programs. Most of the 22 States were considering marginal changes in their eligibility standards.

In the midst of this uncertainty over economic conditions and Federal policy changes, several states were considering added eligibility. During 1981, two states expanded their coverage of pregnant women or infants, and another was considering the coverage of additional low-income children.

A majority of states now extend Medicaid coverage to the medically needy; the options available to states in this area have increased significantly as a result of the amendments included in the Omnibus Budget Reconciliation Act of 1981. In spite of this flexibility, a recent analysis of state Medicaid program plans found that changes in eligibility for and coverage of the medically needy ranked 13 in a list of 15 priority areas of possible change.[16]

Since passage of the 1981 amendments, every state has had to confront the human side of changes in eligibility for its Medicaid program. Early news accounts of state decisionmaking indicate that changes in eligibility standards will continue to be controversial. The structure of the 1981 Medicaid amendments, however, virtually assures that some changes will be made.

The Federal share in almost every state Medicaid program will decline during the next three fiscal years, and state Medicaid costs continue to rise: In May 1981, 26 state Medicaid programs were facing budget deficits. Cost reductions must be found.

Benefits. Responses to the October 1981 survey by the Intergovernmental Health Policy Project (IHPP) indicated that 32 States had adopted or were considering reductions in services, increases in copayments or both (see Table 6). Many of the service changes uncovered by the survey were permitted under Medicaid law prior to the 1981 amendments.

[16] National Governors' Association, forthcoming report on Medicaid.

(1) Twenty-three States were considering the addition of copayment requirements for optional services used by all beneficiaries or for all services used by those eligible for the medically needy program.

(2) Twenty-six States were considering the outright elimination of some optional services provided by their programs.

(3) Twenty-two states were considering further limits on the amount, duration, or scope of mandatory or optional services (e.g., reductions in the maximum number of covered hospital days).

As noted earlier, the National Governors' Association assessment of changes planned as a result of the 1981 amendments found little interest in reducing benefits for the medically needy. That study, however, found a renewed interest in restructuring long-term care services under state programs. The second priority for state implementation of the 1981 amendments was the establishment of community-based long-term care systems.

Prior to the 1981 amendments, Medicaid reimbursement favored institutional services for long-term illness. (By 1978, 42% of Medicaid expenditures went to such institutions; see Table 4.) An amendment included in the Reconciliation Act will allow states to provide a wide variety of community-based home care (excluding room and board) to persons who would otherwise be institutionalized. The services must be provided pursuant to a written plan of care and may not exceed, in the aggregate, what it would have cost to institutionalize the recipient of the care.

It is clear that a great deal of attention will be focused on the service changes that can be made under this provision of the 1981 amendments. It addresses the largest and fastest-growing portion of the Medicaid program and can be implemented in a way that does not cut vital services to patients.

Reimbursement. In October 1981 the largest number of states (37) were looking for ways to change their reimbursement practices in order to control Medicaid expenditures (see Table 6). The IHPP survey found 178 separate reimbursement actions or proposals in those states. Among the most frequent:

(1) Prospective reimbursement for hospitals (or nursing homes). Such proposals require an institution to budget its costs in advance and would establish a Medicaid reimbursement rate based upon the budget.

(2) Reimbursement of hospitals on the basis of an assumed occupancy rate. Hospitals with extremely low occupancy rates must spread their fixed costs over a relatively small number of patient days; the proposals considered in several states would reimburse those hospital costs as if spread over a minimum

number of patient days whether or not the hospital's beds were occupied.

(3) Incentives for outpatient surgery. Several states would reimburse hospitals at only the rate for an outpatient procedure whether or not the patient was admitted.

(4) Reimbursement limits. Several states were considering denial of the annual increase in reimbursement to several types of providers, limiting reimbursement to a fixed percentage of the median of all similar provider costs, reimbursing at emergency room rates only in the case of emergency conditions, or others.

According to the number of items contained in the October 1981 survey report, states were devoting a majority of their time to devising new reimbursement strategies in order to contain the growth of Medicaid programs. The National Governors' Association assessment of priorities for implementing the 1981 amendments confirms the finding: Change in hospital reimbursement methods is the first priority for states.

While changes in reimbursement practices will not have the same direct impact on access to care as denial of eligibility or coverage, it is conceivable that reimbursement levels could be set so low that some hospitals could not afford to treat Medicaid patients. However, a number of states have had experience with prospective hospital budget or rate approval and have been able to implement those programs with a beneficial impact on costs but without adverse impact on access to care.

Summary. In mid-1981, 26 state Medicaid programs were spending above their budgets. To bring budgets into required balance, states were considering changes in eligibility, services, or reimbursement methods. Most states were trying to control their budgets through reimbursement changes, but 22 had adopted or were considering changes in eligibility.

Since the enactment of the 1981 Medicaid amendments, states have adopted reform of hospital reimbursement methods as their first priority. Making changes in eligibility and services for their optional medically needy programs ranks 13th out of 15 priorities.

1983: The Debate Continues

Continued concern about government budgets will keep Medicaid (and other entitlement programs) high on the political agenda for the foreseeable future. The debate over the access of low-income people to medical care will clearly be guided by hospital reimbursement concerns and other implications of the 1981 amendments. Three additional items are likely to approach center stage: copayments for Medicaid services, family responsibility for medical care costs, and nationalization of the

Medicaid program. The remainder of this paper will be devoted to a description of those issues.

Copayments by the patient for a medical care service at the time the service is provided have two basic purposes. First, they are designed to reduce the cost to the third-party payer covering the service. Savings to the insurer either reduce program costs or are devoted to the coverage of additional services. Second, copayments are designed to reduce the utilization of the service to which they are applied. If the patient is required to pay something out of pocket each time he or she sees a physician, trivial visits theoretically will be discouraged. There is conflicting evidence on the disincentive to overutilization at low copayment levels; some observers argue that copayment requirements for low-income people (e.g., Medicaid-eligibles) act as a barrier—not a disincentive—to the receipt of medical care. The administration's FY 1983 budget includes savings to the Medicaid program that will result from requiring states to impose nominal copayments on physician and hospital services used by Medicaid-eligibles.

In recent years, an effort has been made through public policy to prevent the disguise or disposal of assets simply to gain eligibility for Medicaid. Several states have begun to look beyond the assets of the individual to those of responsible family members in assessing the individual's ability to pay for medical care.[17] Critics of "family responsibility" proposals argue that the social insurance philosophy guiding public policy in the United States since 1935 will not countenance the imposition of this kind of a burden on the relatives of the poor and sick. The critics argue that it was the discontinuities of private charity that convinced us, as a society, to replace local poor laws with the Social Security Act. The administration's budget proposal is also based upon changes to Federal law that will permit states to place liens on the homes of nursing home patients and to require contributions from their adult children in order to meet the costs of care.

---

[17] For a discussion of a recent family responsibility proposal, see: Joanne P. Acford, *Reducing Medicaid Expenditures Through Family Responsibility: Critique of a Recent Proposal*, American Journal of Law and Medicine 5:1 (Spring 1979), pp. 59-79. The author concludes that a Massachusetts proposal that would have required financial support from adult children for nursing home care for their indigent parents would weaken family ties and threaten the health care of the elderly. "The plan may have a detrimental psychological impact on the elderly and their adult children, and will have an inequitable economic impact on the children. Furthermore, the imposition of a new legal duty on adult children to contribute to the cost of care for their Medicaid recipient parents contravenes the structure of the health care system" (p. 77).

Finally, President Reagan has proposed to assume responsibility for the full costs of Medicaid at the national level in exchange for state assumption of the full costs of AFDC and food stamps. In terms of access and equity, the implementation of this proposal can have a major impact:

(1) Adoption of the existing national poverty standard would create significant dislocations. Based on 1980 data, 21.2 million ineligible persons below the standard would have become eligible, and 10.4 million Medicaid recipients above the standard would be removed from the program.

(2) Assumption of existing state-defined programs would have no immediate impact. Everyone presently eligible would remain so.

(3) Federalization of the broadest state program would imply prohibitive costs and would still leave millions of persons below the low-income standard ineligible.

(4) National application of the most restrictive state program would remove many presently eligible persons from the program.

In any case, the proposal to "federalize" Medicaid will precipitate a major debate over equity and access. Professor Gutmann, cited at the beginning of this paper, suggests "a sufficiently high level of public provision of health care for all citizens" is one part of the solution to the problem of equal access in a democracy.[18] Debate of her suggestion does not appear on the public agenda in the foreseeable future.

[18] Gutmann, *loc. cit.*, p. 558.

Access to Private Physicians For Public Patients: Participation in Medicaid and Medicare

Janet B. Mitchell, Ph.D.*
Jerry Cromwell, Ph.D.*

The Medicare-Medicaid legislation passed by Congress in 1965 had as a principal goal the incorporation of the elderly and poor into "mainstream" medicine. This goal included improved access to care in private physicians' offices, thereby reducing the reliance on institutional sources of care, where continuity of patient care was often lacking. Yet, since the inception of the two programs, each has been marred by limited physician participation and alleged provider fraud and abuse. These two problems are related, as areas with a high degree of nonparticipation leave the poor to be treated by a small number of physicians with disproportionately high Medicaid caseloads. The pejorative term "Medicaid mills" has often been used to stereotype such practices, in which (it is often assumed) low quality care is provided.

Has access to medical care for the poor been achieved at the expense of quality, or is some care better than no care (or institutional care)? Before attempting to answer these questions we need to know why some physicians choose not to participate in Medicare/Medicaid in the first place and why others choose to see large numbers of these patients. What policy levers can be used to influence the participation decision? What are the credentials and *modus operandi* of these physicians? Is the Medicaid market a lucrative one, or are physicians penalized for treating the poor? In this paper, we examine these questions and similar ones in the area of Medicare assignment.

Ironically, both the Medicaid and Medicare Part B programs have certain features that promote unequal access and physician inequities in treatment of the elderly and the poor. By

* Center for Health Economics Research, Chestnut Hill, Massachusetts.

January 1982

law, Medicare payments to physicians can be no higher than those allowed by private insurers; and Medicaid fees, in turn, must be no greater than Medicare rates. The result is a distinct fee hierarchy with public program rates as much as 50% below usual physician charges. As the lowest-paying program, Medicaid must compete not only with Blue Shield and commercial patients for scarce physician time, but with Medicare patients as well. The elderly poor, furthermore, are at a disadvantage vis-à-vis the more well-to-do elderly. The physician's ability to choose whether to accept Medicare's allowed fee as payment-in-full (i.e., accept assignment) or to bill the Medicare patient the usual fee has essentially created two public insurance programs for the elderly: one for Medicare patients able and willing to pay the coinsurance plus whatever else the physician charges, and another for assigned patients who cannot pay more than the 20% coinsurance.

In addition, public program reimbursement can involve extensive payment delays, prior authorization, limited coverage, and outright administrative denials of claims, all of which raise the collection costs to physicians of accepting Medicare patients on assignment or treating Medicaid recipients. Congressional insistence upon lower reimbursement for public beneficiaries and limits on utilization derived from a natural concern over abuses by patients and physicians alike. The result unfortunately has been the creation of an unattractive secondary market for physician labor, one in which remuneration is lower and the hassles of collecting payment greater. Competition for better-paying, private patients will inevitably drive less qualified physicians into this secondary market in disproportionate numbers.

In the following section of this paper, we discuss how many physicians participate in Medicaid, who they are, where they are located, and the factors influencing their participation decisions. Emphasis is placed on those Medicaid program variables that are particularly amenable to policy manipulation, such as reimbursement and administrative procedures. In the next section we consider similar issues for physicians accepting Medicare patients on assignment. Finally, we discuss the implications of these public program participation rates for the access of the poor and elderly to physician services; and we consider recommendations for public policy.

Medicaid Participation

Variations by Specialty. If the Medicaid program is to achieve its objective of integrating the poor into mainstream medicine, it is essential that private office-based physicians be willing to participate. Although the majority of such physicians do treat Medicaid patients, a sizeable number do not; almost

one out of every four physicians (or 23% nationwide) report never accepting Medicaid (see Table 1).[1] The average participation rate (including nonparticipants) is 12.7%; that is, the average U.S. physician devotes 12.7% of his patient load to Medicaid patients. Most participating physicians have fairly small Medicaid practices, less than 20% of patient load.

TABLE 1: Medicaid Participation by Specialty

Specialty	Average Medicaid Participation Rate (%)	Physicians Who Treat No Medicaid Patients (%)
Primary Care	13.3	21.6
General Practice	13.5	24.3
General Surgery	14.3	8.4
Internal Medicine	14.5	18.1
Obstetrics/Gynecology	8.3	36.8
Pediatrics	14.3	24.1
Medical Specialties	9.0	32.2
Allergy	6.1	40.0
Cardiology	6.7	39.2
Dermatology	13.1	26.1
Gastroenterology	10.0	15.2
Surgical Specialties	13.3	15.3
Neurosurgery	10.9	18.3
Ophthalmology	14.4	12.4
Orthopedic Surgery	10.7	19.8
Otolaryngology	16.2	13.2
Urology	14.3	14.1
Psychiatry	8.0	39.9
ALL	12.7	22.6

Source: Mitchell, Janet B. and Jerry Cromwell, *Medicaid Mills: Fact or Fiction*, HEALTH CARE FINANCING REVIEW 1:37-49, Summer 1980.

As seen in Table 1, there is striking variation in these rates by specialty. Well over one-third of all obstetrician-gynecologists (OB-GYNs), allergists, cardiologists, and psychiatrists decline to participate; and Medicaid patients undoubtedly have difficulty finding available physicians in these specialties. With this notable exception of OB-GYNs, primary care practitioners are more likely to treat Medicaid patients than are subspecialists. General practitioners (GPs), general surgeons, internists,

[1] These data are drawn from Janet B. Mitchell and Jerry Cromwell, *Medicaid Mills: Fact or Fiction*, HEALTH CARE FINANCING REVIEW 1:37-49, Summer 1980. The 15 specialties presented in the table represent 93% of all office-based, patient care physicians in the U.S., excluding anesthesiologists, pathologists, and radiologists. (American Medical Association, PHYSICIAN DISTRIBUTION AND MEDICAL LICENSURE IN THE U.S., 1978. Monroe, WI: AMA, 1979.)

and pediatricians all have above-average participation rates, 13-14%. Compared with their medical specialist colleagues, furthermore, surgical subspecialists also are much more likely to accept Medicaid; in fact, three of the surgical subspecialties have above-average participation levels: ophthalmologists, otolaryngologists (ear, nose, and throat specialists), and urologists.

A good deal of this variation in Medicaid participation rates reflects the demand for particular specialty services by Medicaid beneficiaries. Children and the elderly generate over one-half of total Medicaid expenditures,[2] so it is not surprising to observe high participation rates for specialties treating those age groups. Pediatricians are an obvious example, but otolaryngologists also devote a large proportion of their practice to otitis media, tonsillitis, and other common childhood diseases.[3] Similarly, the elderly constitute a large proportion of the patient population served by ophthalmologists and urologists.[4]

Increased specialist demand due to case-mix differences can only be a partial explanation, however; women of childbearing age represent about a fifth of all Medicaid eligibles, yet OB-GYNs have one of the lowest participation rates of any specialty. In some states such as California, the number of obstetricians accepting Medicaid is reputed to be falling even as the number of Medicaid-eligible women is growing.[5] In physician-scarce areas, such as rural Mississippi, physician nonparticipation forces pregnant women using Medicaid to bypass the local hospital and travel an hour or more to deliver their babies.[6] Yet access to obstetrical care is especially important, because it affects the health of not only the mother but also the unborn child. Why are OB-GYNs so unwilling to treat Medicaid

---

[2] Medicaid/Medicare Management Institute. DATA ON THE MEDICAID PROGRAM: ELIGIBILITY, SERVICES, EXPENDITURES. Baltimore: DHEW, Health Care Financing Administration, 1979.

[3] See National Center for Health Statistics (NCHS), *Office Visits to Otolaryngologists: National Ambulatory Medical Care Survey, United States, 1975-1976*, ADVANCE DATA No. 34, August 30, 1978.

[4] See NCHS, *Office Visits to Ophthalmologists: National Ambulatory Medical Care Survey, United States*, 1975, ADVANCE DATA No. 31, July 14, 1978; NCHS, *Office Visits to Urologists: National Ambulatory Medical Care Survey, United States, 1975-1976*, ADVANCE DATA No. 39, September 7, 1978.

[5] Comptroller General of the United States, REPORT TO THE CONGRESS: BETTER MANAGEMENT AND MORE RESOURCES NEEDED TO STRENGTHEN FEDERAL EFFORTS TO IMPROVE PREGNANCY OUTCOME. U.S. General Accounting Office, HRD-80-24, January 21, 1980.

[6] See, for example, Bleich, Jack, *Some Women Angry over Baby Delivery Situation*, and *Obstetrical Care Still Remains Out of Reach*, THE CLARION-LEDGER (Mississippi), October 30, 1978; Bryant, Barry, *Alabama Ranks Second in Death of Infants*, THE ANNISTON (Alabama) STAR, June 3, 1979.

patients? The most frequently cited reasons are low reimbursement rates and payment delays, but physicians in all specialties have these complaints about Medicaid. However, OB-GYNs are reimbursed after delivery on a package basis that includes prenatal care; thus, Medicaid payment delays could have a disproportionately large adverse effect on cash flow in their practices. National survey data, furthermore, show that nonparticipating OB-GYNs more frequently report these factors as reasons for their nonparticipation.[7] They were not any more likely to cite other reasons, such as a dislike for Medicaid patients. (We will discuss the factors influencing the physician's participation decision in more detail later.)

The very low participation rate for cardiologists is also of particular concern. Given the high prevalence of hypertension among black and elderly patients,[8] we would have expected Medicaid demand for cardiologists' services to be relatively strong. Who then is caring for the majority of Medicaid recipients with hypertension and other cardiac problems? We do not know, but presumably less specialized physicians, such as internists and GPs.

Geographic Variations. Table 2 illustrates striking differences in Medicaid participation, both across geographic regions and between urban and rural areas.[9] Physicians practicing in the West are most willing to treat Medicaid patients, and those in the South and Northeast are least willing. About one-third of physicians in these two regions do not accept *any* Medicaid patients. These differences are largely the result of socioeconomic and Medicaid program characteristics that also

---

[7] Almost two-thirds (61%) of nonparticipating OB-GYNs cited low fees as a very important reason for not seeing Medicaid patients compared with only 44% of nonparticipating GPs, general surgeons, internists, and pediatricians. Similarly, 59% of OB-GYNs said payment delay was very important vs. 45% of the other primary care physicians. For more detail see Mitchell, Janet B. and Jerry Cromwell, *Why Don't OB-GYNs See Medicaid Patients*, working paper prepared for the Health Care Financing Administration.

[8] Medicaid patients include a disproportionate number of blacks. Nonwhites represent 13.3% of the U.S. population, but 35.8% of all Medicaid recipients (Wilensky, Gail, *et al.*, *The Changing Medicaid Population*, paper presented at the American Statistical Association annual meeting, August 1980). Hypertension is the second most common diagnosis for black patients treated in physicians' offices, and the most common for elderly patients. See NCHS, *Office Visits By Black Patients, National Ambulatory Medical Care Survey: U.S., 1975-1976*, ADVANCE DATA No. 50, July 23, 1979; and *Office Visits by Persons Aged 65 and Over: National Ambulatory Medical Care Survey, United States, 1975*, ADVANCE DATA No. 22, March 22, 1978.

[9] This table is based on unpublished data from the national physician survey conducted in 1978 by the Health Care Financing Administration and analyzed by the authors.

TABLE 2: Geographic Variations in Medicaid Participation

Region	Average Medicaid Participation Rate (%)	Physicians Who Treat No Medicaid Patients (%)
Northeast	9.0	30.2
North Central	10.2	20.0
South	8.8	34.0
West	15.6	16.7
Large SMSAs*	9.5	34.6
Small SMSAs	9.8	22.7
Non-SMSAs	14.3	15.2

*Standard Metropolitan Statistical Areas with a population of 1.5 million or more.

Source: Work done by Janet B. Mitchell and Jerry Cromwell under Grant No. 18-P-97723/1-01 for the Health Care Financing Administration, DHHS.

TABLE 3: Credentials of Medicaid Participants and Nonparticipants

Primary Care Physicians	Non-participants (%)	All Medicaid Participants (%)	LMP* Physicians (%)
General Practitioners	46.0	39.9	39.8
Board Certification	31.3	39.5	26.4
Foreign Medical Graduate	10.9	13.8	20.1
Sixty Years or Older	41.5	30.4	36.7
No Hospital Affiliation	5.2	1.8	2.9

* Large Medicaid practices at least 30% Medicaid. These are a subset of all Medicaid participants.

Source: Mitchell, Janet B. and Jerry Cromwell, *Large Medicaid Practices and Medicaid Mills*, JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION 244:2433-2437, November 28, 1980.

vary across regions. The poor are not distributed equally across the U.S., nor are physicians; the South has a relatively larger number of poor persons, for example, but fewer physicians to care for them. State Medicaid agencies also vary widely in how they choose to administer their programs; i.e., how generous they are in reimbursing physicians and in defining Medicaid eligibility. State programs in the Northeast, for example, tend to have a larger Medicaid population because of more liberal eligibility policies, but offer physicians low fees. In the South, on the other hand, Medicaid programs tend to employ very restrictive eligibility criteria, but set comparatively high reimburement rates.

Physician participation in Medicaid generally falls with increasing urbanization; average participation rates are about 50% higher in rural areas than in Standard Metropolitan Statistical Areas (SMSAs). Particularly striking is that one-third of physicians practicing in large SMSAs (defined as those with populations of 1.5 million or more) do not participate at all. In large, physician-dense cities, some physicians may rely on their colleagues to assume responsibility for treating the poor, while in rural areas physicians may participate for charitable reasons, knowing that the poor have few alternative sources of care. Residential segregation of the poor, as in many black urban ghettos, may also contribute to these high nonparticipation rates.

Who Treats Medicaid Patients? Are Medicaid participants different from other physicians? As we have already seen, physicians who participate in Medicaid are more likely to live in a non-SMSA or in the West and to belong to particular specialties. They are also likely to be younger, particularly under 40 years of age.[10] New physicians, anxious to build up their caseloads, may be more willing to accept Medicaid patients. At least in part because of their age, these physicians tend to hold more liberal political attitudes. They are more likely to believe that health care is a right.[11]

Policymakers have been concerned that the better trained, more qualified physicians may have opted out of the Medicaid program. Specialists are actually more likely than GPs to participate, although certain specialties are definitely underrepresented. Board-certification rates, furthermore, are as high if not higher for physicians who treat Medicaid patients.[12] (See Table 3.) Medicaid participants, however, are significantly more likely to have been trained in a foreign medical school, particularly in a Third World (non-Western European, non-English-speaking) country.[13] The average Third World foreign medical

[10] See for example, Hadley, Jack, *Physician Participation in Medicaid: Evidence from California*, Health Services Research 14:266-80, Winter 1979; and Mitchell, Janet B., Jerry Cromwell, and Rachel Schurman, Physician Participation in Public Programs (Chapters 4 and 5), report submitted to the Health Care Financing Administration, September 1981.

[11] Sloan, Frank A., Jerry Cromwell, and Janet B. Mitchell, Private Physicians and Public Programs. Lexington, MA: D.C. Heath, 1978.

[12] See, for example, Sloan, Frank, Janet B. Mitchell, and Jerry Cromwell, *Physician Participation in State Medicaid Programs*, Journal of Human Resources 13:211-245, Supplement 1978; Mitchell, Cromwell, and Schurman, *op cit.*, Chapter 7; Mitchell, Janet B., and Jerry Cromwell, *Private Psychiatrists and Medicaid Participation*, American Journal of Psychiatry, forthcoming 1982.

[13] Studnicki, James, *et al.*, *Foreign Medical Graduates and Maryland Medicaid*, New England Journal of Medicine 294:1153-1157, May 20, 1976; Sloan, Mitchell, and Cromwell, *op. cit.*; Mitchell, Cromwell, and Schurman, *op. cit.*; Mitchell and Cromwell, *Private Psychiatrists*, *op. cit.*

graduate (FMG) devotes almost 50% more of his practice to Medicaid patients than does the U.S.-trained physician.

Because they are disproportionately specialists, Medicaid physician participants are more likely to have admitting privileges at a hospital. This has positive implications for continuity of care, as physicians without such affiliations must either refer patients requiring hospitalization to another physician or send them to an emergency room.

Black physicians are much more likely to treat Medicaid patients than are white physicians, even when specialty and other differences are held constant.[14] Black primary care physicians report an average Medicaid caseload of 29.3%, about two-and-one-half times the size of that of their white colleagues. Of course, black physicians are more apt to treat black patients, who in turn are more likely to be Medicaid recipients;[15] nevertheless, the differences in physician participation are striking.

Limited Versus Full Participation. In Table 1 above and the preceding section, a Medicaid participant was defined as a physician who reports treating any Medicaid patients, even one a week. Some physicians may occasionally treat Medicaid patients in an emergency or for charitable reasons, but otherwise do not participate in the Medicaid program. Thus, some studies have selected a minimum, say ten patients a quarter, to qualify a physician as a Medicaid participant. No matter how Medicaid participation is defined, however, the characteristics of participating physicians are generally the same.

Some physicians may currently treat Medicaid patients but be unwilling to accept any new ones. These physicians have been termed "limited participants" as compared with full participants, who accept all new patients.[16] (Of course, in some cases, a practice may be full, and that is why the physician is not accepting any new patients, Medicaid or otherwise; but here we are concerned with physicians who are taking new private patients but not Medicaid patients.) In a survey of primary care physicians, almost one-half (48%) of Medicaid participants were willing to add all new Medicaid patients into their practice.[17] Only 18% reported that they accepted all new non-Medicaid patients but limited their acceptance of new

[14] Mitchell, Cromwell, and Schurman, *op. cit.*, Chapters 7 and 9.

[15] The vast majority (87%) of all visits to black physicians are made by black patients. See Koba Associates, The Treatment Practices of Black Physicians. Final report prepared for the Office of Health Resources Opportunity, HRA, DHEW, April 1979.

[16] Davidson, Stephen M. and Janet D. Perloff, *Factors Affecting Physician Participation in Medicaid*, paper presented at the American Public Health Association annual meeting, October 22, 1980.

[17] These data come from the national physician survey sponsored by the Health Care Financing Administration in 1976 and analyzed by the authors.

Medicaid patients. Full participants have been shown to be disproportionately located in rural areas and in the Northeast.[18] Their Medicaid caseloads also tend to be much higher than those of limited participants;[19] in this regard they resemble large Medicaid practices, a group we discuss in detail below.

Patient Distribution Across Physician Practices. We know that physicians vary considerably in their willingness to accept Medicaid, but what is the distribution of Medicaid patients across these practices? Are Medicaid patients distributed fairly equally or do they tend to be concentrated within a few physician practices? Here we must account not only for the Medicaid share of the physician's practice, but the size of his caseload as well.

In Table 4, primary care physicians have been ranked from lowest to highest based on the number of Medicaid patients they treat, and then grouped into deciles.[20] The share of total Medicaid patients treated by the physicians in each decile is also shown. If Medicaid patients were perfectly distributed across practices, each decile of physicians would care for 10% of the Medicaid population. In fact, Table 4 indicates considerable unevenness in the distribution of Medicaid patients. Four-fifths of physicians care for only 34% of the total Medicaid population. Almost half (45.6%) of all Medicaid recipients treated in private office practices, on the other hand, are treated by only 10% of the physicians. A small number of physicians appear to have assumed responsibility for a large proportion of the nation's poor. Policymakers have been concerned that these practices with disproportionately high Medicaid caseloads may be "mills" in which low quality care is provided.

Large Medicaid Practices: Medicaid Mills? The Moss investigations spotlighted the problem of Medicaid provider fraud and abuse and, if they were not responsible for coining the phrase, they certainly popularized the term "Medicaid mills."[21] Based on incognito visits in a few large cities, Senator Moss and his staff found the Medicaid-dominated facilities to

---

[18] See, for example, Davidson and Perloff, *op. cit.*; and Mathematica Policy Research, THE PHYSICIAN CAPACITY UTILIZATION SURVEYS: SPECIAL ANALYSES; and table reproduced on page 52 of HEALTH CARE IN A CONTEXT OF CIVIL RIGHTS (Institute of Medicine, Washington, D.C.: National Academy Press, April 1981). The unpublished data cited in note 17 also confirmed their location.

[19] Davidson and Perloff, *op. cit.*

[20] This table is drawn from data presented in Mitchell, Janet B. and Jerry Cromwell, *Medicaid Mills: Fact or Fiction*, HEALTH CARE FINANCING REVIEW 2:37-49, Summer 1980. Similar distributions were obtained with medical and surgical specialists; see Mitchell, Cromwell, and Schurman, *op. cit.*

[21] U.S. Senate Special Committee on Aging, FRAUD AND ABUSE AMONG PRACTITIONERS PARTICIPATING IN THE MEDICAID PROGRAM. Staff report prepared for the Subcommittee on Long-Term Care. Washington, D.C.: Government Printing Office, 1976.

TABLE 4: Distribution of Medicaid Patients Across Physicians

Physicians (%)	All Medicaid Patients (%)	
	Frequency	Cumulative Frequency
0-10	0	0
11-20	0	0
21-30	0.5	0.5
31-40	1.8	2.3
41-50	3.9	4.2
51-60	6.0	12.2
61-70	9.0	21.2
71-80	13.2	34.4
81-90	20.0	54.4
91-100	45.6	100.0

Source: Work done by Janet B. Mitchell and Jerry Cromwell under Contract No. 500-78-0051 for the Health Care Financing Administration, DHHS.

be generally unsanitary, unresponsive to patients, and riddled with unnecessary ancillary services and illegal kickbacks. They characterized these practices as high-volume, assembly-line operations run by well-paid but poorly trained physicians. As a result of these and other investigations,[22] large Medicaid practices (LMPs) and Medicaid mills have generally been considered synonomous, rendering LMPs automatically suspect in the eyes of many public officials.

Analyses based on national survey data, however, show no evidence that LMPs as a group can be characterized as Medicaid mills.[23] LMPs are defined as those in which 30% or more of

[22] See, for example, Bellin, L. E. and F. Kavaler, *Medicaid Practitioner Abuses and Excuses versus Counterstrategy of the New York City Health Department*, American Journal of Public Health 61:200-210, November 1971; Bloom, Lawrence S., *et al.*, *Medicaid in Cook County: Present Status and Future Prospects*, Inquiry 5:13-23, June 1968; U.S. Senate Committee on Finance, Medicare and Medicaid: Problems, Issues and Alternatives. Washington, D.C.: Government Printing Office, February 9, 1970.

[23] These data were based on national surveys of private practice, office-based physicians. It is possible that a particular type of large Medicaid practice was systematically excluded by the sampling criteria; e.g., the inner-city clinic or shared health facility typified by Senator Moss as a mill in his investigations. Thus, we certainly would not want to assert that Medicaid mills do not exist, but rather that private, office-based physicians with large Medicaid practices do not appear to be running Medicaid mills. See Mitchell, Janet B. and Jerry Cromwell, *Large Medicaid Practices and Medicaid Mills*, Journal of the American Medical Association 244:2433-2437, November 28, 1980; Cromwell, Jerry and Janet B. Mitchell, *An Economic Model of Large Medicaid Practices*, Health Services Research, forthcoming.

the physician's patients receive Medicaid; 15.8% of all primary care practices are LMPs. LMP physicians do have higher caseloads than physicians with small Medicaid practices, but they do not appear to hurry patients through their offices; visits are shorter, but only by a minute or two. These high volumes of visits without substantial reductions in the patient's time with the physician are achieved through the greater use of auxiliary personnel in LMPs. More nurses may be used to substitute for physician time, and more clerks to handle the added administrative work associated with Medicaid reimbursement.

There is no evidence that physicians who see large numbers of Medicaid patients are financially abusing the program. There does not appear to be any widespread abuse of ancillary services, such as laboratory tests and x-rays, or excessive markups over costs. LMP physicians at best earn what other physicians make; usually earnings are less. This is true even after adjusting for work effort, training, credentials, and geographic cost-of-living differences.

There is, however, a small group of LMP physicians who do manage to earn very high incomes; one out of every five LMP physicians earns over $80,000 (in 1977 dollars).[24] High incomes in this group are particularly surprising, given low Medicaid fee schedules. How do these physicians manage to earn such high incomes? First, they are more likely to be specialists; but they also see many more patients per week and spend less time with each one. Visit lengths are about five minutes shorter than in small Medicaid practices.

Contrary to the stereotype, most LMPs are not found in large, inner-city ghettos, but rather in small cities, in rural areas, and in the West. As we saw earlier (Table 2), these areas generally also have more participants. LMP physicians, furthermore, tend to be found in solo, unincorporated practices, rather than in large groups.

A serious "credentials gap" does exist, however, between LMP physicians and those operating smaller Medicaid practices. They are significantly less likely to be board-certified, 26% compared with 43% of other participating physicians. (See Table 3.) Almost twice as many foreign medical graduates are found in LMPs than in small Medicaid practices; one out of every five physicians with a large Medicaid practice was trained outside the U.S. Physicians with even larger Medicaid practices (50% or more) are also more likely to be general practitioners and to be at least 60 years old. It thus appears that the Medicaid program and its beneficiaries is a residual market that is secondary to the better insured private market, and that it attracts a disproportionate number of less qualified physicians.

[24] See Cromwell, Jerry and Janet B. Mitchell, *High Income Medicaid Practices*, INQUIRY 18:18-27, Spring 1981.

Surprisingly, this "credentials gap" is a phenomenon only of the Medicaid market for primary care physician services. Medical specialists, surgical specialists, and psychiatrists with LMPs are as well-trained as their colleagues with small Medicaid practices.[25]

Factors Influencing Physician Participation. When asked, physicians consistently cite the same reasons for not participating or for not expanding their Medicaid caseloads. Across the nation, in Medicaid programs as diverse as those of California and Mississippi, low reimbursement rates, payment delays, and other forms of administrative red tape head the list.[26] See, for example, Table 5. Although these physician responses are subjective, statistical analyses have confirmed the importance of such Medicaid programmatic factors in the physician's participation decision. These findings are also remarkably consistent across specialties; from pediatricians to psychiatrists, from allergists to urologists, physicians are more willing to treat Medicaid patients when payment levels are higher and bureaucratic obstacles lower. Physicians also cite more idiosyncratic reasons for not participating; e.g., a general dislike for Medicaid patients or a philosophical opposition to government-financed medical care. We discuss all of these reasons below but emphasize those most amenable to policy manipulation.

Fee schedules. Higher Medicaid fees definitely encourage physicians to treat Medicaid patients; they induce nonparticipants to enter the Medicaid market and encourage participating physicians to expand their Medicaid caseloads.[27] It has been estimated that a 10% increase in the average Medicaid fee would raise participation as much as 3-7% above current levels.[28] Raising Medicaid fees, furthermore, does not appear to foster the development of large Medicaid practices. Evidently, the incentive to expand Medicaid caseloads with more generous fees is offset by the entrance of previous nonparticipants, leaving the *mix* of small and large Medicaid practices unaffected.[29]

---

[25] Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 5; Mitchell and Cromwell, *Private Psychiatrists*, *op. cit.*

[26] See, for example, Jones, M. W. and B. Hamburger, *A Survey of Physician Participation in and Dissatisfaction with the Medi-Cal Program*, THE WESTERN JOURNAL OF MEDICINE 124:75-83, January 1976; Garner, Dewey D., *et al.*, *Factors Affecting Physician Participation in a State Medicaid Program*, MEDICAL CARE 17:43-58, January 1979; Davidson and Perloff, *op. cit.*; Sloan, Cromwell, and Mitchell, *op. cit.*; Texas Medicine, *Why Texas Physicians Are Dropping Out of the Medicaid Program*, pages 359-360 in Allen D. Spiegel, ed., THE MEDICAID EXPERIENCE. Germantown, MD: Aspen Systems, 1979.

[27] Hadley, *op. cit.*; Sloan, Mitchell, and Cromwell, *op. cit.*; Davidson and Perloff, *op. cit.*; Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 7.

[28] Sloan, Mitchell, and Cromwell, *op. cit.*; Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 7.

[29] Cromwell and Mitchell, *Economic Model*, *op. cit.*

TABLE 5: Why Physicians Do Not Participate in Medicaid*

Reason	Reason Considered		
	Very Important (%)	Somewhat Important (%)	Not very Important (%)
Fee schedules too low.	44.4	22.3	33.3
Delays in receiving payments.	40.4	20.4	39.2
It takes too long to fill out forms.	39.6	21.7	38.7
Forms are returned too often for additional information.	38.5	18.3	43.2
Very few Medicaid eligibles in my area.	11.0	12.0	77.0
I'm opposed to government in medicine.	38.8	17.6	43.6
I don't like to treat the kinds of patients who are generally enrolled in the Medicaid program.	12.6	15.1	72.3
I don't have enough time.	26.8	14.2	59.0

*Includes only physicians not participating in state Medicaid programs.

Source: Sloan, Frank A., Jerry Cromwell, and Janet B. Mitchell, PRIVATE PHYSICIANS AND PUBLIC PROGRAMS. Lexington, MA: D.C. Heath, 1978, 122.

Unfortunately, such projections are *ceteris paribus* in the economist's jargon, or assuming nothing else changes. In fact, when one insurer raises payment levels, experience has shown that other insurers will raise theirs as well, thereby eroding any initial gain in participation and fueling physician fee inflation over the long run. There is definite competition among insurers, both public and private, for scarce physician time. When private insurers, such as Blue Shield, offer higher fees, physicians are much less willing to accept Medicaid.[30] Increases in third-party reimbursement rates also encourage physicians to raise their private fees;[31] the net impact of a Medicaid fee increase on access thus might be small or even negative in the long run.

Program administration. Delays in receiving payment also deter participation.[32] Physicians report that Medicaid is far

[30] Sloan, Mitchell, and Cromwell, *op. cit.*; Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 7.

[31] Hadley, *op. cit.*; Hadley, Jack and Robert Lee, *Toward a Physician Payment Policy: Evidence from the Economic Stabilization Program*, POLICY SCIENCES 10:105-120, 1979.

[32] Sloan, Cromwell, and Mitchell, *op. cit.*; Davidson and Perloff, *op. cit.*

more delinquent in reimbursing them than any other insurer: about 3.7 months on average, a full month longer than Blue Shield or Medicare.[33] Where state Medicaid programs have contracted with an outside fiscal agent for claim-processing, participation rates are significantly higher, suggesting that claims may be processed more quickly by private firms.[34] Physicians also complain about the paperwork associated with obtaining reimbursement: time-consuming claim forms, forms frequently returned for more work, retroactive denial of services, etc. These collection difficulties raise the physician's cost of participating in the Medicaid program.

Although state Medicaid programs are mandated to provide physician services, they may impose limitations on the quantity of services, or they may require that physicians receive prior authorization before supplying certain types of services to patients. Both quantity limits and prior authorization directly restrict the supply of physician visits under Medicaid. Prior authorization also may involve considerable administrative red tape. State programs that impose such restrictions definitely pay the price in terms of lower participation rates.[35]

These quantity restrictions may affect two specialties disproportionately: OB-GYNs and psychiatrists. Several states now restrict reimbursement of abortions to only unusual circumstances, such as when the mother's life is endangered, and this number will probably grow over the next few years. Seven state programs have singled out psychiatrists' services for special limitations, either by restricting quantity (one or two visits per year) or by limiting total Medicaid expenditures (e.g., $300 per recipient per year). The District of Columbia's Medicaid program, furthermore, will reimburse outpatient mental health services only within a clinic setting; office-based psychiatrists must obtain prior authorization to treat Medicaid patients.

Medicaid eligibility. Not surprisingly, Medicaid participation rates also vary positively with the relative number of recipients in the physician's community.[36] State Medicaid programs vary in the income criteria used to establish Medicaid eligibility and in whether or not they extend coverage to the "medically indigent." When states are more generous, extending a wider range of services to a larger proportion of the poor, physi-

---

[33] Sloan, Cromwell, and Mitchell, *op. cit.*, 97.

[34] Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 7; Cromwell, Jerry and Janet B. Mitchell, THE ORIGINS AND IMPACT OF MEDICAID REIMBURSEMENT AND ELIGIBILITY POLICIES ON PHYSICIAN PARTICIPATION. Final report submitted to the Health Care Financing Administration, October 1981.

[35] Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 7; Davidson and Perloff, *op. cit.*

[36] Davidson and Perloff, *op. cit.*; Hadley, *op. cit.*; Sloan, Cromwell, and Mitchell, *op. cit.*; Mitchell, Cromwell, and Schurman, *op. cit.*

cians obviously have a larger pool of potential Medicaid patients on which to draw. Not only are average participation levels higher in these states, but we also observe relatively more large Medicaid practices.[37]

Unlike other health insurance programs, Medicaid coverage may fluctuate throughout the year as a function of the individual's eligibility for welfare. The National Center for Health Services Research estimates, for example, that only 43% of the total Medicaid population are actually covered by Medicaid throughout the year.[38] Because of slowness in communication between the central welfare office and the local offices responsible for determining eligibility, physicians may find claims denied retroactively for services delivered in good faith. When Medicaid eligibility lapses, physicians may refuse to provide follow-up treatment, thereby jeopardizing continuity of care. These fluctuations in eligibility pose particular problems for patients with chronic illness and for children in Early Prevention, Screening, Detection, and Treatment (EPSDT) programs. Children may be screened but not treated due to loss of Medicaid coverage, only to return months later to the physician with the same problem.[39]

Practice costs. Physicians also complain that the low fees paid by Medicaid are insufficient to cover overhead expenses. Together with payment delays and uncertainty over payment, some physicians worry that if they see too many Medicaid patients they will be unable to meet payroll and other practice costs.[40] In areas with higher wage rates for nurses and other practice staff, physicians are indeed less likely to participate in Medicaid.[41] This finding has the discouraging implication that wage and price inflation will eventually lead to declining participation rates.

Physician competition. It has been a long-standing policy objective to improve access to care by increasing the supply of physicians. The ensuing competition for patients was expected to encourage physicians to treat historically underserved populations. The effect on Medicaid access, however, appears to be very different from what was intended. In physician-dense areas, physicians are actually much *less* willing to treat Medicaid patients.[42] Those physicians who do participate are more likely to run large Medicaid practices.[43] Thus, competition

---

[37] Cromwell and Mitchell, *Economic Model*, *op. cit.*
[38] Wilensky *et al.*, *op. cit.*
[39] Children's Defense Fund, EPSDT: Does It Spell Health Care for Children? Washington, D.C.: Washington Research Project, 1977.
[40] Garner *et al.*, *op. cit.*; Jones and Hamburger, *op. cit.*
[41] Davidson and Perloff, *op. cit.*; Hadley *op. cit.*; Sloan, Cromwell, and Mitchell, *op. cit.*
[42] Hadley, *op. cit.*; Sloan, Mitchell, and Cromwell, *op. cit.*
[43] Cromwell and Mitchell, *Economic Model*, *op. cit.*

seems to have the perverse effect of further segmenting the public and private markets. And, as we have already seen, it is the less well-trained physicians who are competitively forced to enter the Medicaid market.

These perverse effects hold true only for primary care physicians. Where there are relatively more medical and surgical specialists and more psychiatrists, those specialists are more likely to participate in Medicaid.[44] More specialists are drawn into the program, but with no resulting tendency to concentrate on Medicaid patients. And, as we saw earlier, specialists who do operate LMPs tend to be as well-trained as those with smaller Medicaid practices.

Personal and political attitudes. Of course, personal and political attitudes also play a role in the Medicaid participation decision. Over a third of nonparticipants, for example, cited an opposition to government in medicine as a very important reason in their decision not to accept Medicaid patients.[45] (See Table 5.) Similarly, physicians who strongly disagree that health care is a right have participation levels about 26% lower than those who strongly agree.[46]

Other physicians cite a dislike for the kinds of patients covered by Medicaid.[47] In some instances, this may be due to racism, although there is no good empirical evidence to document this.[48] For the most part, these physicians complain that Medicaid patients frequently break appointments, do not follow instructions regarding medications and other treatment, and tend to "doctor-shop." Nevertheless, physicians who decline to participate or who limit their participation because of a personal dislike for Medicaid patients are a very small minority of all physicians (less than 3%).[49]

Medicare Assignment

The Medicare program, unlike Medicaid, enjoyed relatively widespread acceptance among physicians in its early years. In the late 1960s assignment rates were quite high, approximately

---

[44] Mitchell, Cromwell, and Schurman, *op. cit.*, Chapter 7.
[45] Sloan, Cromwell, and Mitchell, *op. cit.*
[46] *Ibid.*
[47] Davidson and Perloff, *op. cit.*; Garner *et al.*, *op. cit.*; Jones and Hamburger, *op. cit.*; Sloan, Cromwell, and Mitchell, *op. cit.*
[48] Institute of Medicine, *op. cit.*
[49] Sloan, Cromwell, and Mitchell, *op. cit.* Similarly, only 7% of limited participant pediatricians cited this reason (Davidson and Perloff, *op. cit.*). Some physicians may have felt this reason was socially unacceptable and replied that it was not a factor. This of course would lead to an underestimate of its importance.

60% nationwide.[50] Over the last decade, however, a growing number of physicians have been unwilling to accept Medicare patients "on assignment" and thereby limit their charges to Medicare's maximum allowed payment. By 1976, only one of every two Medicare claims was taken on assignment.[51] The willingness of physicians to take assignment has clear implications for the financial liability of Medicare beneficiaries. For nonassigned claims, the patient is responsible for the difference between the physician's actual charge and the Medicare fee (in addition to the 20% coinsurance). The steady decline in assignment rates means that the elderly are bearing an increasingly larger share of their total medical care bill. For the very ill, total liability may be considerable. Low assignment rates may also restrict access of the elderly to physicians' services, especially to ambulatory care that may be critical in avoiding costly hospital and nursing home admissions.

Variations by Specialty. Assignment rates refer to the percent of all Medicare claims taken on assignment and thus are not directly comparable with Medicaid participation rates. A more comparable measure is the percent of physicians' caseloads devoted to assigned patients. As we see in Table 6, about 8-12% of the caseloads of primary care physicians are assigned patients.

Excluding hospital-based physicians who generally have no discretion over claims assignment, general surgeons are more likely to accept assignment than are other specialist groups (see Table 6).[52] Given their generally large bills, surgeons may prefer the certainty of payment associated with assignment. On the other hand, almost one-third of internists and GPs report that they *never* accept assignment.[53] Yet these two specialties are the major sources of care for Medicare patients, accounting for two-thirds of all office visits by the elderly.[54]

Geographic Variations. Assignment rates vary considerably throughout the country, from a low of 34% in the Pacific

---

[50] Waldhauser, Charles B., *Assignment Rates for Supplementary Medical Insurance Claims, January 1968-December 1970*, HEALTH INSURANCE STATISTICS. DHEW Pub. No. (SSA) 72-11702, January 10, 1972.

[51] Ferry, Thomas P., *et al.*, *Physician Charges Under Medicare: Assignment Rates and Beneficiary Liability*, HEALTH CARE FINANCING REVIEW 1:49-73, Winter 1980.

[52] *Ibid.* The data presented on assignment rates refer to the percent of services taken on assignment. Assignment rates based on the percent of charges taken on assignment tend to be somewhat higher, especially for surgical subspecialists.

[53] Mitchell, Janet B. and Jerry Cromwell, PHYSICIAN BEHAVIOR UNDER THE MEDICARE ASSIGNMENT OPTION. Final report submitted to the Health Care Financing Administration, January 1981.

[54] NCHS, ADVANCE DATA No. 22, *op. cit.*

TABLE 6: Medicare Assignment by Specialty

Specialty	Average Assignment Rate (%)	Physicians Who Never Take Assignment(%)*	Caseload Assigned (%)*
General Surgery	49.9	21.1	12.0
Family Practice	48.6		
Cardiology	47.3		
General Practice	46.7	32.0	7.6
Orthopedic Surgery	46.0		
Urology	45.7		
Dermatology	44.3		
Internal Medicine	44.1	27.7	9.4
Otolaryngology	35.4		
Ophthalmology	35.3		

*Physician and caseload percentages from Mitchell, Janet B. and Jerry Cromwell, PHYSICIAN BEHAVIOR UNDER THE MEDICARE ASSIGNMENT OPTION. Final Report submitted to the Health Care Financing Administration, January 1981. Data are based on physician self-reports.

Source: Average assignment rates from Ferry, Thomas P., *et al.*, *Physicians' Charges under Medicare: Assignment Rates and Beneficiary Liability*, HEALTH CARE FINANCING REVIEW 1:49-73, Winter 1980. Data are based on Medicare claims.

Northwest to a high of 67% in New England.[55] In Table 7, a national sample of GPs, general surgeons, and internists have been grouped according to whether they always, never, or sometimes, (i.e., on a case-by-case basis) take Medicare assignments.[56] Physicians in the Northeast are far more likely to accept at least some claims on assignment. Perhaps most striking is the very small number (9.1%) of Western physicians willing to assign *all* of their claims, less than half the number in other parts of the country.

Just as they are more likely to accept Medicaid patients, physicians in nonmetropolitan areas are also more willing to accept assignment than are their city colleagues. This is probably due, at least in part, to the fact that the elderly in rural areas are apt to be poorer and to have more limited supplementary insurance coverage. The physician is more likely to accept assignment when he feels the patient could not pay the bill.[57]

[55] Ferry *et al.*, *op. cit.*

[56] Mitchell and Cromwell, MEDICARE ASSIGNMENT, *op. cit.*

[57] Wakely, J. T., *et al.*, A STUDY TO DETERMINE FACTORS PHYSICIANS TAKE INTO CONSIDERATION IN DECIDING WHETHER TO TAKE ASSIGNMENT, BILLING ARRANGEMENTS PHYSICIANS MAKE TO COPE WITH THE DEDUCTIBLE AND CO-INSURANCE FEATURES OF REIMBURSEMENT, AND THE NATURE OF THESE BILLING PRACTICES. Vol. 1. Final Report submitted to SSA by Research Triangle Institute, August, 1972.

TABLE 7: Geographic Variations in Willingness To Accept Medicare Assignment

	Never (%)	Sometimes (%)	Always (%)
All Physicians	28.2	52.9	18.9
Northeast	19.0	56.6	24.4
North Central	35.1	47.6	17.3
South	29.3	50.0	20.7
West	32.3	58.6	9.1
Large SMSA*	30.7	55.3	14.0
Small SMSA	28.8	51.7	19.5
Non-SMSA	24.4	50.8	24.8

*Standard Metropolitan Statistical Areas with a population of 1.5 million or more.

Source: Mitchell, Janet B. and Jerry Cromwell, PHYSICIAN BEHAVIOR UNDER THE MEDICARE ASSIGNMENT OPTION. Final report submitted to the Health Care Financing Administration, January 1981.

Assignment is mandatory, furthermore, for those elderly who are also Medicaid-eligible.[58] Thus, we also observe higher assignment rates for very old and for black beneficiaries, characteristics associated with limited ability to pay and Medicaid eligibility.[59]

Factors Influencing Medicare Assignment. Physicians who take Medicare assignment are more likely to be surgeons and to live in non-SMSAs, particularly in the Northeast. Like Medicaid participants, they are also more likely to be younger and to be more liberal in their political outlook.[60] In contrast to the Medicaid program, physicians treating assigned Medicare patients are as well-trained as those who do not. There are no differences in board-certification,[61] and although foreign medical graduates are more likely to accept assignment, their average caseloads are not any larger than those of U.S. graduates.[62]

[58] In the case of patients eligible for both Medicare and Medicaid, the physician is reimbursed the full allowed Medicare charge, not the lower Medicaid fee. The state Medicaid program is then responsible for paying the deductible and 20% coinsurance not covered by Medicare. (Four state programs, Alaska, Louisiana, Oregon, and Wyoming, have chosen not to "buy" into the Medicare program.)

[59] Ferry *et al.*, *op. cit.*

[60] Mitchell and Cromwell, MEDICARE ASSIGNMENT, *op. cit.*; Paringer, Lynn, *The Medicare Assignment Rates of Physicians: Their Responses to Changes in Reimbursement Policy*, HEALTH CARE FINANCING REVIEW 1:75-89, Winter 1980.

[61] Mitchell and Cromwell, MEDICARE ASSIGNMENT, *op. cit.*

[62] Mitchell and Cromwell, MEDICARE ASSIGNMENT, *op. cit.*; Paringer, *op. cit.*

Physicians who participate in Medicaid are also more likely to accept Medicare assignment and to supply more assigned visits *in toto* than nonparticipants.[63] Since Medicaid fees are lower than Medicare fees, physicians will not begin to see Medicaid patients until they have exhausted the available pool of Medicare-assigned patients.

Medicare reimbursement and administration. Unlike those of Medicaid, Medicare reimbursement policies are determined nationally, ensuring some degree of uniformity across states. Nevertheless, administrative responsibility for the actual reimbursement of services is delegated to carriers that serve designated geographic areas. As a result, there is considerable variation in average payments and in billing and collection procedures. When Medicare reimbursement levels are higher, physicians are definitely willing to accept more of their Medicare patients on assignment and to devote a larger share of their practice to assigned patients.[64] Similarly, assignment rates are higher when the administrative burden of obtaining reimbursement is lower; e.g., when fewer Medicare claims are investigated or are denied altogether.[65]

Raising allowable Medicare fees in order to encourage assignments may have unexpected side effects. First, the physician may add *Medicare*-assigned patients to his caseload, but at the expense of lower-paying *Medicaid* patients. Second, assigned patients, because of the increased coinsurance, must now pay more for the same visit; and thus access may be restricted for the poorer of them. Finally, raising Medicare-allowed charges will also stimulate increases in physicians' private fees because Medicare-eligibles can collect more from the program, thereby attenuating at least some of the gain in assignment.[66] (We discuss these issues in detail later.)

Medicare eligibility. Assignment rates are actually lower in areas with relatively greater proportions of Medicare bene-

---

[63] Paringer, Lynn, The Medicare Assignment Rates of Physicians: Their Responses to Changes in Reimbursement Policy. Urban Institute paper, March 1979.

[64] See, for example, Huang, Lien-Fu and Orest Koropecky, The Effects of the Medicare Method of Reimbursement on Physicians' Fees and Beneficiaries Utilization, Volume 2, Part 1. Washington, D.C.: Robert R. Nathan Associates, April 1973; Institute of Medicine, Medicare-Medicaid Reimbursement Policies. Washington, D.C.: Government Printing Office, March 1, 1976; Mitchell and Cromwell, Medicare Assignment, *op. cit.*; Paringer, 1980, *op. cit.*

[65] Mitchell and Cromwell, Medicare Assignment, *op. cit.*; Muller, Charlotte and Jonah Otelsberg, *Carrier Discretionary Practices and Physician Payment under Medicare Part B: A Preliminary Report*, Medical Care 17:650-666, June 1979.

[56] Paringer, 1980, *op. cit.*; Hadley and Lee, *op. cit.*

ficiaries.[67] This seems paradoxical, but the assigned and nonassigned components of the Medicare program definitely compete against each other for scarce physician time. In communities with a disproportionately large elderly population, the physician has many more nonassigned patients he can treat before beginning to accept poorer patients on assignment. Serious barriers to access may arise in the future as a growing elderly population competes for physician services.

Physician competition. Physician competition appears to have the same unexpected effect on Medicare assignment that it does on Medicaid participation. In physician-dense areas, physicians have lower assignment rates and smaller assigned caseloads.[68] It is possible that these communities are characterized not only by fewer participants, but also by the concentration of large numbers of assigned patients within a few physician practices (as with our large Medicaid practices); to date, however, there is no evidence concerning Medicare maldistribution.[69]

Summary and Policy Implications

Economics Count. One clear implication of the evidence regarding physican participation in public benefits programs is that economics count—medical need alone is insufficient in explaining how physicians distribute themselves, choose to perform certain procedures, and decide not to treat Medicare- and Medicaid-eligibles. A medical education costs so much today, in both terms of tuition and books and personal sacrifice, that physicians must pay some attention to insurance coverage and reimbursement schedules. They are also in such great demand that most can be selective in taking on new patients, putting less attractive (or lucrative) ones further back in the queue used to fill up their (or their staff's) "down" time.

Unfortunately, those relegated to the back of the line, or to the less qualified physician, are not randomly distributed but generally are poorer and less well insured. Medicaid-eligibles, and to a lesser extent Medicare enrollees, are often less well insured, from the physician's perspective, because of the lower fees received. Additional red tape associated with payment only reinforces his/her financial prejudice against treating the elderly and the poor.

---

[67] Huang and Koropecky, *op. cit.*; Mitchell and Cromwell, MEDICARE ASSIGNMENT, *op. cit.*

[68] Mitchell and Cromwell, MEDICARE ASSIGNMENT, *op. cit.*; Paringer, *op. cit.*

[69] Senator Moss writes of the existence of Medicare mills in Florida, but they have not been investigated. See Moss, Frank E., *Through the Medicaid Mills*, in Spiegel, *op. cit.*, 387-392.

Drawbacks to Nonparticipation. Physician nonparticipation has many undesirable outcomes. First, the absolute number of physicians available to the poor is less than optimal, as one-quarter refuse to treat any Medicaid patients or to take Medicare beneficiaries on assignment.

Second, limited participation results in more competition among the elderly and the poor for scarce physician time and narrower choices in terms of proximity and quality. Public patients in general have difficulty finding a private physician in larger cities—especially an obstetrician-gynecologist or medical specialist. Medicaid patients, in particular, experience access problems throughout the South, while fewer physicians take the elderly on assignment in the West.

Third, while both public programs clearly improve overall access, the lower fees bias participation in favor of the less-skilled physician, creating a *credentials gap* between the public and private programs.

Fourth, it is inevitable in such an atmosphere that Medicaid mills occasionally arise, protected by fee differentials from competition from more qualified physicians.

And fifth, lower assignment rates among Medicare physicians place an additional financial burden on their elderly patients who must pay the higher private fees.

Raising Medicare/Medicaid Fees. The path to better access for public program beneficiaries would appear straightforward, but as we shall see, it is not. Raising Medicaid or Medicare fees will substantially increase physician participation/assignment rates, but is this the optimal program strategy from society's point of view? Insurer competition poses a serious dilemma in this regard. Because Medicare assignment and Medicaid compete with each other, as well as with the nonassigned component of the Medicare program and with private insurers, any reimbursement change in one program will have downstream consequences for the others.

Among the current public insurance programs, Medicare nonassignment offers the most generous reimbursement, because the physician simply bills patients at his usual rate. The physician naturally will treat all Medicare patients willing and able to pay his usual fee before accepting any patients on assignment. Since Medicaid fees by law must be no higher than Medicare's maximum allowed charge, the physician will then see as many assigned patients as possible before he begins to treat non-elderly Medicaid patients (at least in meaningful numbers). Furthermore, public policy permits differential treatment of the elderly who are Medicaid-eligible. The physician must accept these patients on assignment, but he is reimbursed

at the relatively higher Medicare rate.[70] Thus, the elderly poor enjoy somewhat better access to physician services than the younger poor.

What exactly happens if we raise public reimbursement levels? An increase in the Medicaid fee schedule would induce Medicaid physician participants to expand their Medicaid caseloads. As Medicaid fees still may not exceed the Medicare charge (and thus encourage the substitution of Medicaid for Medicare patients), Medicare assignment rates would remain unchanged, and total public program participation for these physicians would necessarily increase. Some of the increase undoubtedly would come at the expense of the noninsured poor. Many of these individuals, in fact, might be former Medicaid recipients who have lost their eligibility; of the almost 12 million recipients who were Medicaid-eligible for part of 1977, 55% were uninsured the rest of the time.[71]

Because of patient copayments and the fact that physicians can accept assignment on a case-by-case basis, the impact of a Medicare reimbursement increase is much less clear. With higher Medicare-allowable fees, some physicians would be attracted into the assignment market for the first time, while others would expand their Medicare caseloads, including their supply of assigned services. All physicians treating the elderly, however, would enjoy a windfall gain as a result of the higher rates (including those physicians who never accept assignment), because the patient/Medicare would be willing and able to pay more. In effect, an income transfer would be made from the government (and Medicare patients) to physicians in an effort to induce them to devote more time and resources to the elderly.

Gains in access, unfortunately, would not accrue to all Medicare enrollees. The relative position of nonassigned patients would improve, as they would pay less for the same visit because they now could collect more from Medicare. Ironically, this would be the wealthier elderly group that presumably would have experienced little difficulty obtaining access to care heretofore. Meanwhile, assigned patients would be actually worse off, as they would pay *more* for the same visit because of the 20% copayment on a higher allowable charge. This would place a particularly onerous burden on the nonwhite elderly who are much less likely to have supplementary insurance coverage;[72] some might even forego care as a result.

---

[70] Except in the four state programs that have not "bought" into Medicare, Medicaid pays the deductible and coinsurance (based on Medicare's allowed charge) for joint eligibles.

[71] Wilensky *et al., op. cit.*

[72] Berstein, Amy B. and Marc L. Berk, *Perceived Health Status and Selected Indicators of Access to Care among the Minority Aged*, working paper, National Center for Health Services Research, undated.

Two other considerations further detract from unilateral fee increases as a simple solution to the participation problem. So far, we have ignored the response of private insurers like Blue Shield to Medicare/Medicaid fee increases. In order to assure continued good access for their beneficiairies, they would certainly raise their allowable rates if at all possible. This obviously would dilute the initial Medicare/Medicaid participation impact, leading to a second round of increases. Of course, nothing is inherently wrong with government having to compete in the private market for resources; it happens all the time. The problem here is that very weak competitive limits on fees exist in the private market because of insurance. Most of the competition among insurers (on behalf of their beneficiaries) redounds to the financial benefit of a scarce resource, namely physicians, whose net incomes already average over $63,000.[73]

With such high incomes, it is even possible that many physicians would respond perversely to even higher fees by working less.[74] Insurer competition under this scenario of "backward-bending supply" would be especially frustrating, as access for all kinds of patients, private as well as public, would be restricted.

Recommendations

So long as different public programs (or parts of programs) compete with each other, physicians will discriminate in favor of the patients paying higher fees. Eliminating the assignment option and equalizing reimbursement for Medicare/Medicaid services should eliminate current inequities and ensure equal access for all public beneficiaries. The latter goal is difficult to achieve, however, because of insurer competition. Ideally a single regulatory body would establish a uniform fee schedule for both public and private insurers, but a national health plan of this type is unlikely in the short run. Other, less radical approaches, however, could also help encourage competition among insurers and providers. Such approaches might include: prepaid insurance plans, taxing the employer's contribution

[73] Gaffney, John C. and Gerald L. Glandon, eds. PROFILE OF MEDICAL PRACTICE 1979. Monroe, WI: American Medical Association, 1979.

[74] Whether physicians' supply curves are indeed backward-bending is still a matter of empirical investigation. See, for example, Feld-stein, Martin S., *The Rising Price of Physicians' Services*, THE REVIEW OF ECONOMICS AND STATISTICS 52:121-133, May 1970; Sloan, Frank, *Physician Supply in the Short Run*, INDUSTRIAL AND LABOR RELATIONS REVIEW 28:549-569, 1975; Vahovich, Stephen G., *Physicians' Supply Decision by Specialty: 2SLS Model*, INDUSTRIAL RELATIONS 16:51-60, February 1977.

towards health insurance, the use of vouchers and tax credits to encourage shopping for less costly plans, etc.

Physician participation in public programs could definitely be enhanced by streamlining the billing and collection procedures involved; e.g., reducing payment delays and eliminating unnecessary red tape. Since cost control must inevitably be balanced against improvements in access, the marginal program dollar may be better spent on improving payment mechanisms than on raising fees. This would certainly imply large initial outlays in some states, but these costs would be non-recurring, while fee increases would remain in perpetuity.

Finally, care must be taken in correcting Medicare/Medicaid fraud and abuse not to discourage physicians from treating the poor. Low fees exact a financial penalty on all participating physicians, but blanket criticism of large Medicaid practices may discourage even the most altruistic physicians from expanding their practices in underserved areas.

Does Equal Access Mean Equal Quality? The partitioning of the public market resulting from unequal reimbursement levels means not only that the lowest-paying (Medicaid) patients are the last to be added to physicians' caseloads and the first to be dropped, but also that they are likely to be treated by less-qualified physicians. A primary goal of the Medicaid program to integrate the poor into mainstream medicine is thereby thwarted. Of course, many of the poor under Medicaid are receiving office-based care for the first time and access has thereby improved, but it is by no means equal. Assess to ambulatory physician care, whatever the background of the physician, may be better than no care at all (or institutional care); it may even be adequate; but it is not equal.

But does equal access necessarily mean equal quality? Are the poor entitled to the same quality of care as others? Experience with other public subsidies such as housing suggests not. While minimum standards should be required for safety reasons, most people would probably agree that other measures of "quality" (say, size of living space or number of closets) need not be equal between the public and private sectors. Physicians participating in the Medicaid program of course do meet the minimum standards required by law (e.g., licensure), and there is no evidence to show that the care rendered by non-board-certified or foreign-trained physicians leads to relatively poorer health outcomes. Nevertheless, perceived quality differences can sometimes be as important as real ones.

In the final analysis, we as a society must decide whether equal access includes equal quality (however defined), and if so, whether we are willing to pay the price. Ironically, many of the physicians whom we expect to provide these services to the poor do not believe that the poor have a right to the same

medical care as everyone else.[75] Furthermore, the copayment provisions under Medicare suggest that policymakers have viewed access to physician services as a "limited" right; the elderly are expected to pay some amount out-of-pocket for all visits and to pay more to see higher-quality (more expensive) physicians. The Medicaid program, on the other hand, was implemented with no deductibles or coinsurance, implying that medical care is a right for persons below a certain income level. A number of state Medicaid programs have begun considering the use of copayments for physician services,[76] however, so even this is in doubt. Because of the ambiguities and inequities involved, society must eventually resolve this dilemma of medical care as a right versus a privilege.

[75] A national sample of primary care physicians were asked whether they agreed or disagreed with the following statement: "It is the responsibility of society, through its government, to provide everyone with the best available medical care, whether he can afford it or not. . ." Over one-half of the physicians (53.5%) stated that they disagreed with this statement. See Cromwell, Jerry, Janet B. Mitchell, and Rachel Schurman, *Physician Political Attitudes*, working paper prepared for Health Care Financing Administration.

[76] Spitz, Bruce, STATE GUIDE TO MEDICAID COST CONTAINMENT. Washington, D.C.: Intergovernment Health Policy Project and the Center for Policy Research, National Governors' Association, September 1981.

Equity in the Distribution of Quality of Care

Leon Wyszewianski, Ph.D.*
Avedis Donabedian, M.D., M.P.H.*

The purpose of this paper is to examine the question of whether the quality of health-care services in the United States is equitably distributed across all groups in the population. Until recently that question had received less attention than the closely related question of whether there is equity in the distribution of the *quantity* of health-care services. It is becoming widely recognized that differences in the amount of services used by various groups have been narrowing, to the point that, for example, persons in lower-income groups now make more physician visits than persons in high-income groups (Figure 1). As these changes take place, the quality of services provided to different groups has begun to receive increasing attention.

A few years ago a much-publicized series of investigations by the U.S. Senate Special Committee on Aging, under Senator Frank Moss, sought to document the fraud, abuse, and poor quality of care given by "Medicaid mills."[1] Those investigations revealed the importance of quality as a consideration in measuring the success of government-financed efforts designed to achieve equity in the distribution of health services. Indeed, if the steady increase in the utilization of services by lower-income groups that has been attributed in large part to the Medicare and Medicaid programs only reflects greater use of the kind of inappropriate care Senator Moss was denouncing, then the apparent progress made in reaching an equitable distribution of health services may just be an illusion and a cruel hoax for both the recipients of the care and the taxpayers who pay the bill.

*Department of Medical Care Organization, School of Public Health, The University of Michigan, Ann Arbor, Michigan.

This study, which was originally prepared for the Commission, is reprinted with permission from MEDICAL CARE, Vol. 19, No. 12 (Supplement), 1981.

February 1981

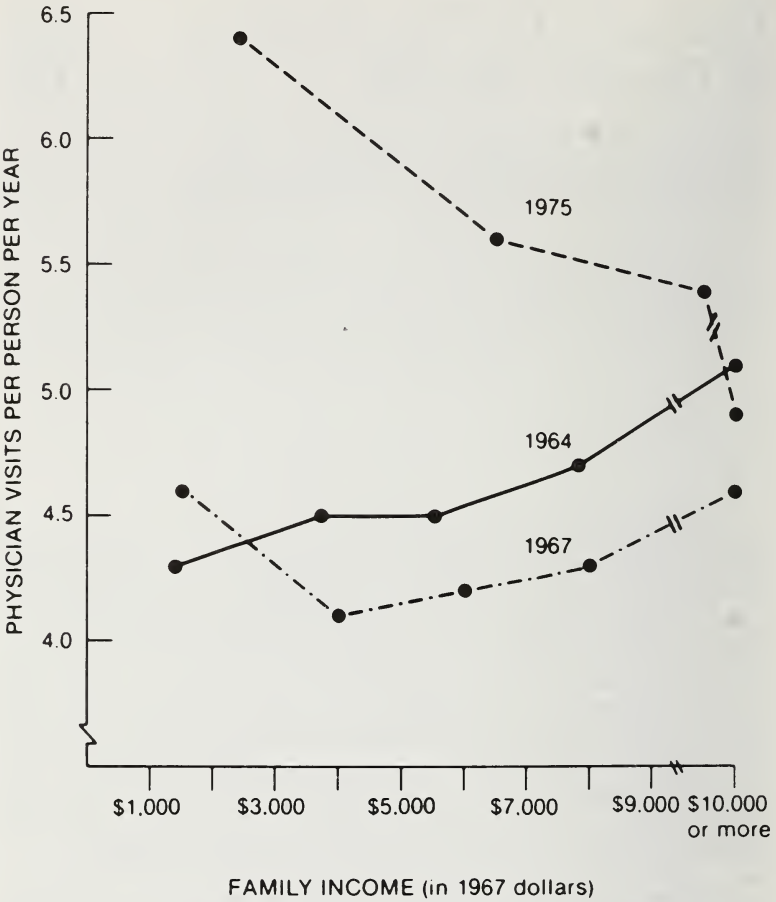


Fig. 1 Number of physician visits per person per year, by family income, adjusted to 1967 dollars (U.S.A., civilian noninstitutionalized population, 1964, 1967, (1975). From Donabedian et al.[66] Reproduced with permission.

Although these and other similar developments in the last few years have brought forward a concern with the equitable distribution of the *quality* of health care services, that concern is actually not new to the field. Almost half a century ago Lee and Jones had already included, in their now classic description of what constitutes good medical care, the stipulation that medical care cannot be considered good unless it is within the reach of everyone:

> Good medical care implies the application of all the necessary services of modern, scientific medicine to the needs of all people. Judged from the viewpoint of society as a whole, the qualitative aspects of medical care cannot be dissociated from the quantitative. No matter what the perfection of technique in the treatment of one individual case, medicine does not fulfill its functions adequately

until the same perfection is within the reach of all individuals.[2]

We begin our examination of equity in the distribution of quality of health care by describing a framework for conducting that examination and by defining and discussing the concepts that underlie the framework. Using this framework to organize our discussion, we then proceed to an examination of the relevant empirical evidence. In a final section we present a brief summary and discuss the implications of our findings for future data collection and research efforts, as well as for the formulation of policy.

Framework and Definitions

General Framework. Equity in the distribution of the quality of health care is examined in this paper within a framework that consists of a series of increasingly stringent tests for the presence of equity in the distribution of health services among the general population. This series of tests and the presumptive inequities to which the tests relate are represented by the set of numbered questions (Figure 2), under the terms of which the framework is described. In subsequent subsections we discuss the conceptions of quality and equity that are implicit in our formulation of the framework:

(1) The first and perhaps crudest measure of whether there is an equitable distribution of health care services among different population groups is to compare the average use of services for the groups of interest. If the averages differ, there are grounds for presuming that structural or access inequities exist among the groups. Thus, by this test, the use of physician services by the different income groups shown in Figure 1 is considered inequitable—to different degrees and in different directions—for all three years shown: 1964, 1967, and 1975.

(2) Recognizing, however, that population groups differ with respect to their need for health services, the above test for inequity can be refined by adding the notion of need, and asking whether there is, for each group, an equal match between services used and the number of services needed. If, as has often been shown, lower income persons tend to experience more illness than those in high income groups, those in lower income groups need more physician services; therefore, the distribution shown in Figure 1 for 1964 is—by this new test—even more inequitable than may have appeared at first. By contrast, the distribution for 1975 may be, using this criterion, closer to reaching equity. This approach to testing for equity is the one followed by Aday et al.,[3] who examined the use of services in relation to need, based on data from the 1975-76 national household survey of health care utilization and expenditures conducted by the Center for Health Administration Studies at the University of Chicago. When they related two

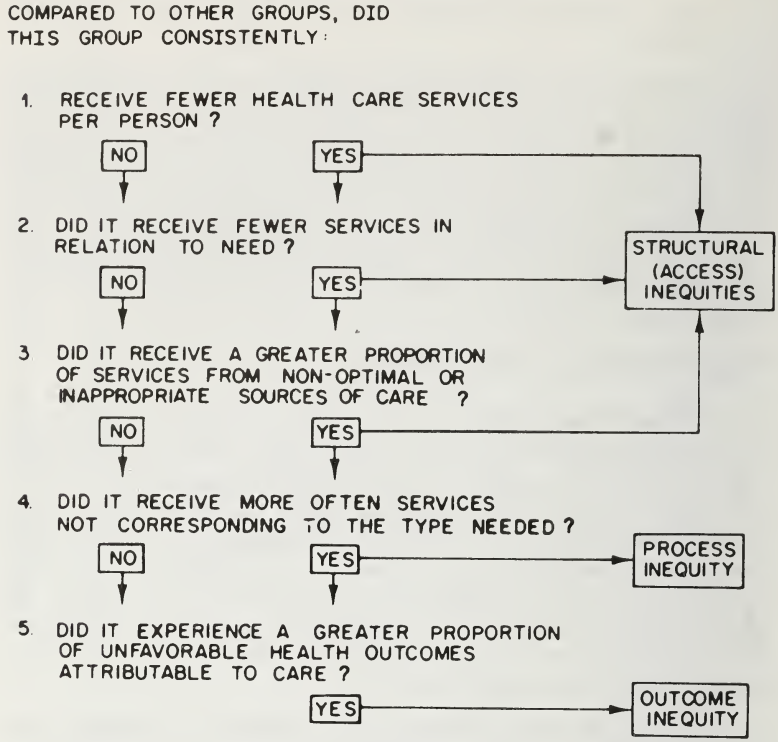

COMPARED TO OTHER GROUPS, DID THIS GROUP CONSISTENTLY:
1. RECEIVE FEWER HEALTH CARE SERVICES PER PERSON ?
NO
YES
2. DID IT RECEIVE FEWER SERVICES IN RELATION TO NEED ?
NO
YES
STRUCTURAL (ACCESS) INEQUITIES
3 DID IT RECEIVE A GREATER PROPORTION OF SERVICES FROM NON-OPTIMAL OR INAPPROPRIATE SOURCES OF CARE ?
NO
YES
4. DID IT RECEIVE MORE OFTEN SERVICES NOT CORRESPONDING TO THE TYPE NEEDED ?
NO
YES
PROCESS INEQUITY
5. DID IT EXPERIENCE A GREATER PROPORTION OF UNFAVORABLE HEALTH OUTCOMES ATTRIBUTABLE TO CARE ?
YES
OUTCOME INEQUITY

Fig. 2. General framework for the classification of inequities in the distribution of health care.

different measures of need to the levels of utilization in 1975-76, they found that, in contrast to data from their earlier surveys, great improvements had been made in the direction of equity among nearly all population groups.

(3) However, as Aday *et al.* note, the recent lessening in the inequities they found in relation to the *quantity* of health services only heightens the importance of going one step further and examining the cost and effectiveness of the services being provided. As we thus proceed to ever more stringent tests of equity, attention shifts from quantity to quality. In Figure 2, the first question asked from the perspective of quality is one that relates, as the first two questions do, to concerns with access to care: did all groups receive the same proportion of services from appropriate or even optimal providers? Ideally, all services used should have been provided by properly qualified health professionals in facilities that have all the necessary equipment and support staff.

One way to put this test into operation might be to examine whether all groups seeing physicians have access to what Payne and his associates[4] term the "modal" specialist, that is, a physi-

cian specifically trained to diagnose and treat the condition for which care is being sought. Payne and his associates have shown that such specialists provide a higher quality of care within their area of specialization than either general practitioners or other specialists providing care outside the domain of their specialty.

(4) However, even if all groups are served in equal proportion by "optimal" providers, it is conceivable that such providers deliver better care to some patients than to others. Therefore, to test for that possibility it is necessary to ask whether all patient groups received, to the same extent, care deemed appropriate for their conditions. Any departure from equal treatment so defined presumes that inequities exist in relation to the process of delivering care.

(5) The final test for inequities focuses on the ultimate results of the care provided to patients. The question asked here is whether all groups exhibit a similar proportion of positive or negative health outcomes attributable to care. As is well known, factors other than care can be responsible for health outcomes and health status. For that reason, if the test for process inequities described above reveals that none are present, we would, ideally, expect no outcome inequities. By contrast, process inequities may not always translate into outcome inequities because two groups can conceivably have similar levels of health status even though one of the groups is given less adequate care.

As reflected in the structure of Figure 2, the different types of inequity are not mutually exclusive. On the contrary, the different types are likely to be present simultaneously, giving rise to what Hart,[5] in an examination of inequities in the distribution of health services in Britain, calls the Inverse Care Law, whereby the availability of high-quality medical care to a population varies inversely with the population's need for care. In the worst case, it is possible to conceive of a group that has higher needs for care than other groups, but receives less care, from less qualified providers, who in turn provide even less adequate care to the members of this group than to other persons. The end result is that the group has much worse outcomes and lower health status levels than other groups.

Within the framework just described, our focus is only on those distributional inequities that relate to quality of care—that is, with issues that correspond to questions 3, 4, and 5 in Figure 2. As we already noted, those issues have received less attention than the issues that correspond to questions 1 and 2, which focus on equity with respect to the *amounts* of services used; these aspects have been examined most recently by Aday et al. using data from the series of surveys conducted by the Center for Health Administration Studies and the National Opinion Research Corporation (CHAS/NORC) at the University of Chicago.[3] The same set of issues was examined earlier by

Andersen and his associates,[6] and was also based on CHAS/NORC data.

Before turning to the empirical evidence relating to the different aspects of equity in the distribution of quality of care, we examine briefly the concepts that underlie the framework described above, beginning with the concept of equity.

Equity. As we define it, for the quality of health care services to be equitably distributed, the services provided must be equal with respect to (1) the appropriateness of the source of service for the condition being treated and (2) the degree to which the services themselves represent the most efficacious approach, according to current medical knowledge, for the condition being treated.

Our definition of equity is analogous to those adopted by Andersen[6] and Aday et al.[3] in their respective examinations of equity in the distribution of the quantity of services. Their definitions and ours not only emphasize equality in services provided, but also equality in the degree of match between services needed and services received. Their definitions of need are formulated in terms of amounts of services; ours is directed at the type of service required. Both kinds of needs are related to one another, as we have mentioned, and we address that point in greater detail in the subsection on the relation between quantity and quality.

Theoretically, any departure from the mean level of service for the population represents a form of inequity whereby some people receive more or better services while others receive fewer or less appropriate ones; in practice, the focus has been on inequities that are associated with the characteristics of certain groups, which is reflected in the statement by Aday et al.:

> Inequity in health service distribution occurs when individuals receive services primarily according to their place in the social structure, their enabling characteristics, or the characteristics of the health system instead of their needs.[3]

Much like Andersen and his associates and Aday et al., our main concern is with inequities that are associated with characteristics such as age, sex, race, income, ethnic background, and area of residence. We are thus interested in the presence of consistently inferior care provided to persons who fall into the kinds of groupings that Aday et al.[3] and Davis and Schoen[7] have identified as least likely to receive the services they need. They include, among others, the aged poor, residents of inner cities, and certain members of minority groups.

We are not so much concerned, therefore, with inequities that arise, for example, from the occasional errors of judgment or technique that a practitioner might make and that result in isolated instances of inferior care. Patients receiving care that is

inferior compared with the care the same practitioner's other patients receive are of interest in this inquiry only if they differ in some special way from the practitioner's other patients. Put another way, equity implies a concern with those differences in treatment which, because they exhibit a systematic pattern, are evidence of discrimination. Such a concern, of course, reflects the belief, expressed by Lee and Jones in the statement quoted earlier, that everyone should have access to equally good medical care.

It should be stressed, however, that equity implies a concern with relative rather than absolute levels of quality. Conceivably, quality of care could be equitably distributed and still be abysmally poor for all. In addition, to establish the presence or absence of equity so defined, it must be possible to draw comparisons among diverse groups; information on one particular population group, no matter how good or how poor the care for that group, is not sufficient to establish whether there is equity or inequity. Thus, in a strict sense, many of the studies that Brook and Williams[8] included in their examination of the quality of care received by the disadvantaged are not, of themselves, evidence of inequity to the extent that those studies focus solely on the groups included under the rubric of disadvantaged, even though those groups often do indeed receive poor quality care.

The Relation Between Quality and Quantity. In examining equity in the distribution of quality of care we exclude explicitly the consideration of equity with respect to amounts of services received. Although valid, the distinction is also in part arbitrary, reflecting our judgment that the consideration of equity with respect to the *quantity* of services has benefited from the availability of pertinent national data and that it has also been the object of study not only by those associated with the CHAS/NORC surveys[3,6] but by a relatively large number of other researchers and scholars, as reflected in Verbrugge's recent review of that line of inquiry.[9]

By contrast, the explicit consideration of issues of equity with respect to quality of care has not received the same level of attention, in part, perhaps, because until some level of equitable distribution of the amount of services is reached, the saliency of the quality of those services will not be great. We are only now beginning to reach, as we already noted, equity in the distribution of quantities of services. A review of the effects of Medicaid and Medicare on access and quality of health care points out:

> Obviously, quantity is a necessary, though not sufficient, prerequisite of quality. To the extent that segments of the population that were formerly deprived have gained access to care, an opportunity has been created for improvements in quality; to the extent that use is still insufficient relative to need, progress toward the attainment of quality is im-

peded. But access to larger quantities of care has hazards of its own. It exposes people to care that may be unnecessary, although harmless. More often, unnecessary care is also associated with a variety of dangers of different magnitudes and probabilities of occurrence.[10]

In a later section we examine some examples of care being inappropriate when more care is being provided to some groups.

The Definition and Measurement of Quality

Structure, process and outcome. Studies of quality of care can be distinguished in terms of whether they evalute "structure," "process," or "outcome":

(1) Appraisal of *structure* involves "the evaluation of the settings and instrumentalities available and used for the provision of care." Studies of structure, then, evaluate quality in terms of the characteristics of the persons and organizations involved in the provision of care.

(2) Assessment of *process* is "the evaluation of activities of physicians and other health professionals in the management of patients."

(3) Assessment of *outcomes* is, in turn, "the evaluation of end results in terms of health and satisfaction."[11,12]

The distinction between structure, process, and outcome is reflected in the framework in Figure 2. Though it is useful to distinguish those three ways of evaluating quality of care, each of them is only meaningful in the context of the others.[13,14] The characteristics of structural elements such as personnel and facilities are of interest for evaluating quality of care only to the extent that these characteristics are associated with how clinical activities are carried out—that is, with how satisfactory the process of care is. The nature of the activities that constitute the process of care are themselves of interest from the perspective of quality only to the extent that they are believed to contribute to desired health outcomes. Health outcomes, in turn, can be taken as a meaningful measure of quality only if they can be tied to specific elements of process or structure; they may otherwise be attributed to other factors not related to care, such as the environment, public health measures, and genetic endowment.

The need to identify the links between structure and process and between process and outcomes is implicit in our framework for identifying inequities in the distribution of quality of care. For example, to establish the existence of such inequities, it is not sufficient to determine that there are differences in the structure of care between two population groups; that is, one is served by one type of facility or provider and the other is served by another type. Such differences are

evidence of inequity with respect to quality only if one type of provider has been shown to be more likely to provide inadequate or inferior care than the other.

Similarly, identifying differences in health outcomes, such as infant mortality rates, is not sufficient to establish that inequities in the quality of care are present. To reach that conclusion it is necessary to demonstrate that differences in outcomes can be attributed to care. Therefore, the interest is not in absolute differences in infant mortality rates among population groups, but in differences expressed in terms of *preventable* deaths; that is, deaths where care was, or could have been, a decisive factor.

Assessment criteria and equity. The evaluation of quality of care, whether in terms of structure, process, or outcomes, requires that criteria be formulated as to what represents health care of good quality. Those criteria ideally reflect current knowledge about the relations between structure and process, and between process and outcome. Unfortunately, our knowledge concerning the relation between process and outcome (the relation between what the practitioner does and its effect on the patient's health status) is itself, in many instances, more tentative than is desirable. Often there is no conclusive proof that a given procedure has the effects generally attributed to it.[15] Even in such instances, however, it is possible to formulate criteria for assessing care, since there is usually a high degree of consensus among practitioners as to what is the most desirable course of action in particular situations, whether or not there is conclusive proof of the efficacy of the procedure.

Different types of criteria, and different approaches to formulating criteria have been identified.[12] Whatever the approach, however, most criteria as to what constitutes good medical care are likely to prove inadequate to evaluate the care provided for the type of person who is especially likely to receive inequitable treatment in terms of quality of care. Such persons include, in particular, those who bear the double burden of poverty and ill-health. Treatment of their conditions, to be successful, may require going beyond strictly clinical intervention, which is not likely to be reflected in the standard criteria to evaluate the medical care process. Yet the implications for quality of care of such special needs was already recognized by Lee and Jones:

> Good medical care is coordinated with social welfare work. "Where love of man is, there also is love of this art." But love of man must be clarified by an understanding of man and must take note of his social environment and economic needs. The interrelations of sickness and poverty have long been recognized. A low standard of living weakens the bodily resistance to infection, and poor ventilation, overcrowding, and malnutrition are strong allies of disease. The value of medical care among the poorer economic classes is vitiated if the patient is sent back after a course of treatment

into the same conditions which gave rise to his ailment. While medicine cannot cure unemployment or abolish slums, much can be done by cooperation with the social agencies to improve the level of health among the poorer members of society.[2]

Even on a more modest scale than that envisioned by Lee and Jones, certain social and economic characteristics of a patient may require that "standard" clinical strategies be modified to take account, for example, of the patient's inability or unwillingness to make a return visit. To the extent that the need for such adjustment is not reflected in process criteria used for evaluating care, we may observe situations in which inequities in quality of care are found when quality is evaluated in terms of *outcomes*, but no such inequities are observed when the usual *process* measures are used. As we shall see, such situations seem to exist, and they raise difficult questions as to what constitutes the proper boundaries of medical care services. The boundaries drawn by Lee and Jones, for example, are broader than what many in the field would consider appropriate.

The technical and interpersonal components of quality. Most quality of care studies focus on the interaction between an individual provider (usually a physician) and an individual client. The emphasis is on how the care affects the client's physical-physiological function, as opposed to its effects on the client's psychological or social functioning.[16] Within the evaluation of that kind of provider-client interaction, two additional components have been distinguished, the "technical" and the "interpersonal" aspects of care:

(1) Technical performance has been defined as "the application of medical knowledge and technology in a manner that maximizes its benefits and minimizes its risks, taking account of the preferences of each patient."[17]

(2) Management of the interpersonal relationship with the client or the "art of care" has been defined by Brook, et al. as relating to the "milieu, manner, and behavior of the provider in delivering care to and communicating with the patient."[18]

These two components—the technical and the interpersonal—have been acknowledged to be of approximately equal importance in evaluating care.[16,18] Yet most quality of care studies have focused on the technical component, in part because "the technical management of health states . . . is the technology we understand best, the application of which is our most distinctive function, and that often makes the most difference in situations of acute peril to life and limb."[19]

More generally, because the technical aspects are those that are best understood and because they are the ones most likely to be documented in the medical record and elsewhere, they are

much easier to measure and evaluate than the interpersonal aspects of care. The latter are seldom documented, and well-defined performance criteria to assess them have yet to be developed.

To the extent that the interpersonal aspect of care is an important component of health services, the relative lack of information on the quality of performance in terms of that component has special implications for any inquiry into the equity of the distribution of quality of care. As Brook and Williams[8] have noted, the groups most likely to suffer from inequities are also those for whom the problems of communication with providers may be the most severe. Brook and Williams cite some studies that illustrate this issue, but they are all surveys of the satisfaction of patients with various aspects of care, including physician-patient communication. Although such reports on patient satisfaction are an indication of the quality of the interpersonal component of care, a fully valid measure of this component must be explicitly related to desired outcomes. Put another way, because the technical and interpersonal components are two aspects of the overall process of care, measures used in evaluating either one of those components are meaningful only if they, like all process measures, have a specified relation to outcomes.

The interpersonal and technical aspects of care can be seen as forming a mutually reinforcing set;[14] there is indirect empirical support for this view in the finding, by Riedel and Riedel, that "physicians who were patient-oriented (i.e., who enjoyed the relationship aspects of their practice) usually received higher scores [on the technical aspects of care] than physicians who preferred to be involved in diagnosis and treatment."[20] Similarly, Payne and associates have made the following case, based on their findings, for the technical and interpersonal aspects of care that have a joint effect on outcomes:

> In urinary tract infection the freedom from infection was related to high physician treatment performance only in patients who observed the medical regimen. . . . This moderator, patient compliance, was the key to relating good physician performance to good outcomes in essential hypertension and in iron deficiency anemia . . . It is apparent that the skills of the physician in diagnosis and therapeutics must be supplemented by communication skills and human relations skills in order to achieve good outcomes through patient cooperation in the medical regimen.[21]

Client and provider perspectives. Just as quality of care has been measured primarily in terms of its technical aspects, and partly because of it, quality of care assessments reflect for the most part the judgment of providers as to what constitutes appropriate care. It is all the more important, therefore, to

recognize the place of the client's perspective in the assessment of care. From the perspective of quality of care, patient satisfaction is itself a desired outcome; it therefore is of interest for that reason if for no other. Patient satisfaction has also been included as a factor in models, such as the health belief model, that are designed to account for a person's initiation of care and the continuation of that care.[22] The seeking of care and the continued use of it, however, are not as directly relevant to quality of care as they are to concerns about access and the quantity of services used. It is not surprising, therefore, that Aday et al.[3] measured patient satisfaction and included it among the factors they examine in their most recent consideration of the equitable distribution of health services.

A recent review of the empirical evidence examining the client's and the provider's view of quality concludes that there are many similarities between the two, with concern about technical competence being central to both. As might be expected, clients place greater weight on the interpersonal aspects of care.[14] The commonalities in the client's and the provider's views of quality notwithstanding, from the perspective of equity in the distribution of quality, the interesting situations are those in which there is divergence between the client's and the provider's view of what constitutes appropriate care. If, for example, the physician gives in to the patient's demands for treatment that the patient considers appropriate but the physician considers unnecessary (such as prescribing tranquilizers or an antibiotic when medically they are not indicated) patient satisfaction is apt to be higher—a positive outcome. On the other hand, the patient's health status may be affected adversely by the unnecessary medication. Weighing such benefits against their associated costs is the kind of difficult calculus that faces every provider. From the perspective of equity, however, the specific questions that must be asked are:

(1) Do providers give in more easily to the demands of some patients than those of others?

(2) Are some providers more prone to give in or even cater to client demands, either out of the conviction that, on balance, it represents the best course of action, or as part of a calculated or even unscrupulous strategy to retain their patients?

Either of these two possible situations represent instances of inequitable treatment because certain types of patients or the patients of certain providers are more likely to experience outcomes that are different from those of other patients. Although the levels of satisfaction of those patients might conceivably be higher, the technical care provided to them is judged inferior because it exposes the patients to at least some risk, and it does not represent a properly parsimonious use of resources, which itself is one of the attributes of good medical care.[14]

The Empirical Evidence

Introduction and Overview. In the preceding section we set out a framework to examine the question of whether there is an equitable distribution of quality of health care in the United States. We also defined some of the concepts and specified the relationships that underly this framework. As we now turn to an examination of the empirical evidence, we begin by describing in general terms the kind of evidence that we see as necessary for our purpose, and indicate to what extent that evidence appears to exist.

The evidence we are seeking relates to questions 3, 4, and 5 in Figure 2, which correspond to the three types of inequity that we identified as structural, process, and outcome inequities. The ideal set of data to answer the questions in Figure 2 consists of information about the type and quality of care provided to a suitably diverse cross section of the United States population, with sufficient representation from groups that are particularly apt, because of their social or geographic situation, to receive less appropriate care than the rest of the population. Put another way, the data we desire represent the equivalent, for *quality* of care, of what already exists, with respect to *quantities* of care consumed, from the National Health Interview Survey and the CHAS/NORC surveys.

Data on quality of care actually available fall far short of those expectations, as has previously been noted by Brook and Williams when they examined the issues of quality of health care for the disadvantaged,[8] and by Donabedian in his examination of the effects of Medicare and Medicaid on access and quality of care.[10] In both papers the authors commented on the paucity and uneveness of pertinent studies on quality of health care.

Because we exclude from our area of concern the issue of equity with respect to the quantity of care used (because others have already dealt with it in some detail) the only studies that are relevant here are those that evaluate quality of care and do so across different population groups. The number of such studies that are reasonably sound methodologically is not great. Unfortunately, several of the studies of quality of health care from the past few years that had a comparatively large scope did not report any results in terms of the kind of patient characteristics that are of interest here such as age, sex, income, or area of residence. Examples of such studies are a major study by Riedel and Riedel[20] as well as one by Payne and associates.[21] Lyons and Payne had reported on variations in quality of care by age of patients in an earlier study,[23,24] but they do not report any data based on patient characterists in their latest study.

Brooks and Williams[8] relied heavily on studies of quality that focused on the "disadvantaged." Because they do not de-

fine that term explicitly, the selection of the studies they cite has some circularity to it: studies of some groups appear to have been included because they showed the group to have poor health status, or because the group appears to have received poor care. In addition, as we noted earlier, studies focusing on a specific "disadvantaged" population are of limited value in this inquiry because a concern with equity must emphasize the relative rather than the absolute levels of quality of care provided to any given group.

Furthermore, most of the evidence presented in this paper is drawn from studies that are concerned with physician care provided either in the hospital or in the physician's office. Our heavy reliance on such studies reflects in part our own greater familiarity with that portion of the quality of care literature. It also reflects the underdeveloped state of quality assessment studies concerned with care provided by persons other than physicians and in settings other than the hospital. For instance, with respect to studies of quality of care in nursing homes—a setting in which there is reason to be concerned about inequities—a review article by Weissert notes that "problems of reliability and generalizability plague much of the research" in that area.[25]

In the next sections we present and discuss the pertinent evidence we were able to identify. We do so in three subsections, each corresponding to one of the types of inequities included in our framework, beginning with structural inequities.

Structural Inequities. In terms of the framework we specified earlier, structural inequities are present whenever certain groups receive a greater proportion of their care from individual providers or health care organizations that are either not optimal or are inappropriate for the needs of the individual persons within the group. Those represent, then, "qualitative access" inequities, as distinguished from the "quantitative access" ones that correspond to questions 1 and 2 in Figure 2. Thus the focus is not on the appropriateness of the amount of services provided, but on whether the source of whatever services were obtained is likely to provide appropriate care in relation to the person's needs.

The presence of structural inequities is more difficult to document than may at first appear. In many instances where differential use of sources of care has been shown to be related to factors such as race or income rather than the nature of the patient's clinical condition, it is not clear that the different sources involved provide care that is significantly different in its quality. Moreover, when differences in quality can be demonstrated, the source of care considered to be more desirable is not necessarily the one that provides the higher quality level.

The study by Egbert and Rothman[26] is a good illustration of striking findings that nevertheless do not constitute conclusive evidence of structural inequities. Egbert and Rothman examined all records of patients who underwent two common surgical procedures, inguinal hernia repair and gallbladder operations, at one hospital from 1952 to 1972. Throughout that period black patients, when compared with white patients, were much more likely to have the operation performed by a surgical resident than a staff surgeon, as shown in the left panel of Figure 3. As the right-hand panel of Figure 3 makes clear, the trend between 1952 and 1962 was in the direction of greater equity between the two groups; between 1962 and 1972 the direction changes toward even greater inequity. According to Egbert and Rothman, the differences found could not be attributed to the patients' emergency status. However, Medicaid patients were found to be significantly more likely to be operated on by staff surgeons than non-Medicaid patients (47% compared with 13%) and no significant differences by race were found among Medicaid patients.

To conclude from those findings that structural inequities were present at the hospital studied by Egbert and Rothman it is necessary to agree with the authors when they state that they "assume that resident surgeons are less skilled than staff surgeons since resident surgeons are in training; the staff surgeons are the teachers of the resident surgeons."[26] The implication is that compared with patients who had operations performed by staff surgeons, patients whose operations were performed by residents have a greater likelihood of having negative outcomes such as surgical complications and surgical mortality.

Though reasonable on the surface, such a set of assumptions is not easily made. Our system for training surgeons and other physicians is based on a quite different assumption, namely, that a well-supervised resident in training provides care that is of similar quality to that usually provided by the supervising physician or physicians; this perspective is reflected in the work of Perkoff and Anderson.[27] In a study of patients who came to the emergency room of Barnes Hospital, a major teaching hospital in St. Louis, Missouri, they found that the assignment of patients to ward or private status was not related to their ability to pay, as might have been expected, but to their race. Even among those who had no private physician, white patients were apparently more likely to be referred to a private physician than black patients. According to Perkoff and Anderson, however, this inequitable treatment may have worked to the advantage of black patients in terms of quality of care:

> It must be pointed out that this situation cannot be assumed always to work to the disadvantage of the Negro. The quality of medical care for the ward patients is ex-

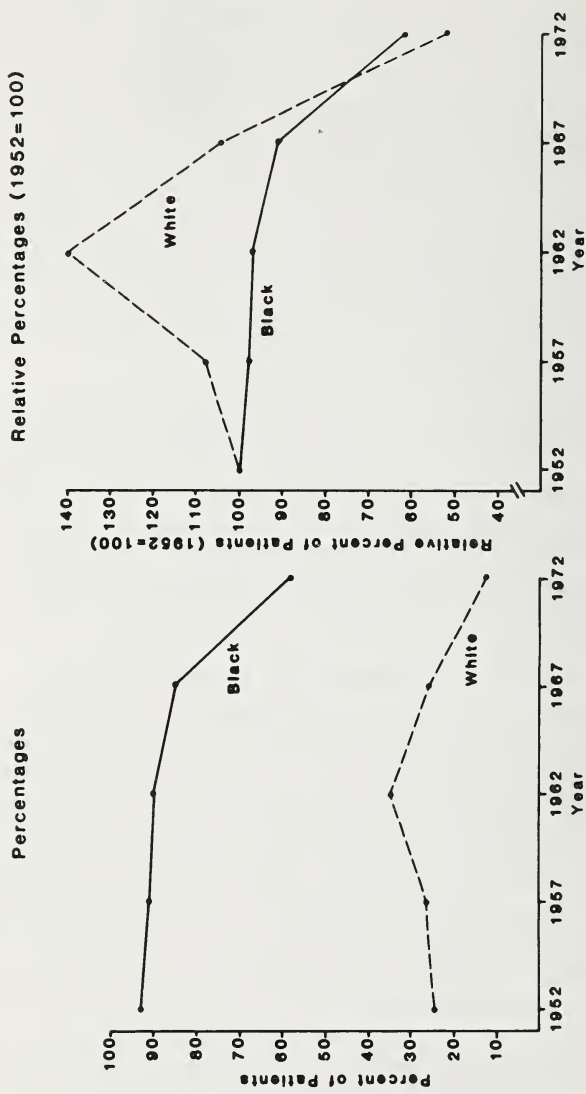


FIG. 3. Percentages and relative percentages of patients of specified race who had inguinal hernia repairs and cholecystectomies performed by the resident staff of a hospital in Baltimore, Maryland, 1952–1972.[26]

> tremely high, with considerably more available expert consultation than is the case for private patients. Indeed, while the active referral of Caucasian patients to private physicians usually provides them with excellent physicians, the care is more variable and less rigorously supervised than is that for ward patients.

On the other hand, contrary evidence is provided by a classic study of perinatal mortality in New York City by Kohl,[28] which showed that ward patients had a greater percentage of preventable perinatal deaths than did private patients, and that house staff physicians, who provided most of the care to ward patients, had a higher frequency of preventable deaths than either obstetricians or other physicians.

We are thus faced with contradictory perspectives and evidence on how structural differences (that is, differences in the source of care) translate into quality of care differences. In terms of our earlier discussion of structure, process, and outcome, we are faced with alternative interpretations of how structural differences are likely to relate to process and outcome elements of care, with little objective basis for choosing among them. Most studies reporting the presence of structural differences that are unrelated to medical need do not provide information on how those differences relate to process or outcomes aspects of care. In that respect, the missing element in the study by Egbert and Rothman is information on surgical management and on surgical outcomes such as rates of surgical deaths and post-surgical complications.

So far our discussion of the relation between structural differences and quality of care has centered on technical performance. We stated earlier that the other aspect of care, the interpersonal component, has received comparatively less attention, in part because it is more difficult to measure. Yet paradoxically, it may be easier to relate the kind of structural differences in care that we have been discussing to the quality of the interpersonal component than was the case for technical care.

Duff and Hollingshead[29] provided, some years ago, a vivid and detailed picture of the differences in the interpersonal aspects of care in three different types of accommodations in a large medical center—private, semiprivate, and ward—to which patients were assigned in ways that reflected their status in the community. As the investigators saw it, the hospital staff's attitudes toward patients depended on two status indicators: (1) whether the patient had a private physician and (2) the type of accommodation that patient was in:

> Although having a personal physician did not always ensure that physician's attention, it did protect the patient against some of the resented practices found in the ward accommodations where patients were more often the tools

> of learning than the objects of concern in the eyes of those who cared for them. . . . On the wards suspicion, distrust, and confusion were extreme and common. The performance of physicians tended to keep alive this distrust. The patient hoped for the best care and an explanation of what was being done. Although he may have been given optimal treatment, if he did not get an explanation of what was being done, he often felt that he was being used as a "guinea pig."[29]

The introduction of Medicaid was expected by many to have the effect of eliminating the kind of structural differences and attendant differences in interpersonal aspects of quality described by Duff and Hollingshead. The evidence on that score is at least mixed. The findings by Egbert and Rothman, cited earlier, indicate that, at least in teaching hospitals, patients who are considered "teaching material" continue to be disproportionately black and are drawn from the lower socioeconomic strata of society. Similarly, a study by Dutton[30] of the patterns of use of care by residents of Washington, D.C. show that, at least in that city, persons using public clinics and hospital outpatient departments as their usual source of care had incomes that were almost one-half those of persons for whom the usual sources of care were solo or group private practitioners, or prepaid group practices. In the clinics and outpatient departments serving these lower-income persons the level of satisfaction of patients with physicians and that of physicians with patients were both lower compared with the other settings; this suggests that quality of care in terms of interpersonal relations was also lower, and that, at least with respect to the interpersonal aspects of care, structural inequities were present.

Such problems in the interpersonal aspects of the care provided to lower income groups may appear less severe, however, by comparison with the fraud and abuse reported in the treatment of Medicaid patients in New York City.[31-35] The more general issue of "Medicaid mills" received wide publicity as a result of the report of the U.S. Senate Special Committee on Aging chaired by Senator Frank Moss.[1] Most of the charges against such "mills" included unnecessary referrals (the "ping-ponging" of patients), overprescribing, ordering of unnecessary tests, and outright fraud by billing for services not delivered. In addition, the "mills" were described as providing low-quality care, which was attributed in large part to their having physicians who are less well trained and less specialized than most physicians in the community.

Mitchell and Cromwell[36] recently set out to determine whether Medicaid mills are a common nationwide phenomenon or whether they are a relatively rare occurrence found primarily in New York and Chicago. Specifically, they sought to establish whether physicians in large Medicaid practices

(LMPs)—practices in which at least 30% of the patients are Medicaid patients—have the characteristics that physicians in Medicaid mills have been described as having to a disproportionate degree: older, more general practitioners, more foreign medical graduates, more without hospital affiliations, and handling much larger caseloads. They conclude, based on their analysis of data from a 1976 physician survey conducted by the National Opinion Research Center, that "it does not appear that most LMPs are Medicaid mills." They note, however, that high-income LMPs (HIPMPs)—those in which physicians reported net incomes of $80,000 or more—included some practices that have mill-like characteristics. HILMP physicians were found to be older, to be more likely to be foreign trained, to have far larger incomes, and to have much shorter visit lengths. Contrary to expectations, however, they were also more specialized.

Even if most LMPs did not conform to Mitchell and Cromwell's criteria for classifying them as Medicaid mills, important differences emerged from a comparison of, on one hand, LMPs and EXLMPs—"extra-large" Medicaid practices having 50% or more Medicaid patients—with, on the other hand, small Medicaid practices (SMPs), which included practices with fewer than 30% Medicaid patients but excluded practices that had no Medicaid patients at all. The findings confirm the general profile of Medicaid practices that had already emerged from the more fragmentary information available a few years ago.[10] Compared with SMP physicians, LMP and EXLMP physicians were found to be significantly less likely to be board-certified. In addition, EXLMP physicians were found to be significantly more likely to be: (1) general practitioners, (2) 60 years of age or older, and (3) less likely to be affiliated with a hospital. All those characteristics repeatedly have been found to be associated with lower quality care, most recently in the work of Riedel and Riedel.[20]

On another aspect of LMPs and EXLMPs, Mitchell and Cromwell state that:

> LMP physicians not only devote a large portion of their practice to Medicaid, but they also see the majority of all Medicaid patients treated in private practices. The EXLMPs alone see 31.8 per cent of all Medicaid recipients, and together with other LMP physicians they provide care to over one-half of this Medicaid pool (58.4 per cent).

On the other hand, only 5.5% of physicians were found to be in EXLMPs, and 15.5% were in LMPs. Thus a large proportion of the total Medicaid population is served by a small percentage of all physicians, and those practitioners have significantly lower qualifications than physicians whose practices include only a small percentage of Medicaid patients. Put another way, the results presented by Mitchell and Cromwell provide strong evidence that structural inequities exist with respect to Medi-

caid patients, even if "true" Medicaid mills are much less prevalent than had been believed. Even on that last point it should be noted that the data Mitchell and Cromwell used were entirely based on information provided by the physicians themselves in response to a telephone survey, and as the authors acknowledge: "Non-participation and non-response among physicians operating true Medicare mills is still likely . . . How serious a bias this creates is unknown."[36]

So far our examination of structural inequities has concentrated on care provided by physicians in and out of hospitals, which reflects, as we noted earlier, the strong bias that exists toward such services in the literature on health services in general and on quality of care in particular. However, as is clear from an earlier summary of reported Medicaid abuses,[10] a wide range of providers and institutions were found to deliver unnecessary or poor quality services to Medicaid recipients: pharmacists, dentists, podiatrists, optometrists, and nursing home operators. To the extent that the qualifications and characteristics of these various providers exhibit patterns similar to those we described for the physicians who provide most of care received by Medicaid patients, they too would constitute instances of structural inequity. Unfortunately the necessary comparative data to make that determination are not readily available.

State-financed institutions are another setting in which structural inequities are very likely to exist. As Brook and Williams[8] have pointed out, such institutions not only are often staffed by less than fully qualified foreign medical graduates, but their populations are primarily drawn from lower income groups. Such is the case for health services in prisons[37] and state mental health hospitals. The general field of mental health has been the subject of classic studies that identified the kind of striking structural inequities that are perhaps less prevalent nowadays. For instance, Myers and Schaffer[38] provided a dramatic illustration of structural and other inequities in their study of a psychiatric clinic. They examined the relation between social class and whether treatment for "neurosis" was recommended, as well as the relation between social class and the type of provider assigned to the patient. They found that no treatment was recommended to more than 67% of the patients in the lowest social class, but only 10% of those from the highest class received the same recommendation. Almost 75% of the lowest class patients were assigned to medical students, and none of them was assigned to a resident or staff psychiatrist. By contrast, over 40% of the highest class patients were assigned to staff psychiatrists, nearly half to resident psychiatrists, and none to medical students. Similar relations between social class and type of treatment of mental illness were reported in another classic study by Hollingshead and Redlich.[39]

Before examining process inequities, by way of summary we point to the main characteristics of the evidence about the existence of structural inequities:

(1) Evidence of differential use of providers that is related to nonmedical characteristics of patients is not necessarily an indication that structural inequities in terms of technical quality exist. However, such differences are likely to be related to structural inequities in terms of interpersonal aspects of care, about which there is more concrete evidence than might be expected.

(2) Many of the available studies focus on a single institution, making generalizations difficult, and, for that reason broader-scale studies are much needed.

(3) Despite the general paucity of data, there is evidence that structural inequities exist. This is particularly true with respect to recipients of Medicaid, a program that had as one of its ostensible goals the elimination of precisely the kind of access-related inequities discussed in this section.

Process Inequities. We now consider inequities that might arise in relation to another aspect of care, the delivery process itself. Whenever certain groups receive services that are judged less appropriate to their needs in contrast to services received by other groups, that, in the framework we have postulated, represents a case of process inequity.

Most assessments of the quality of medical care have been studies of the activities involved in caring for patients—that is, the process of care. That process has been evaluated in a number of ways, most of them variants of two approaches:

(1) The use of expert judges who evaluate the care, usually by reviewing the medical record, basing their judgment on their own, internal criteria;

(2) The use of a list of explicit criteria that represent the activities judged to be appropriate for the care of certain types of patients or conditions. Such lists are often compiled with the participation of at least some of the practitioners whose care is to be evaluated. As we noted earlier, the criteria are usually designed to reflect the best knowledge available as to what actions are most likely to lead to desired health outcomes for specified clinical situations.

Even though there is a relative abundance of quality of care studies using process measures, many of them are not suited to our purpose, either because they focus on small, relatively homogeneous patient populations or because they only report on the relation between quality of care and practitioner characteristics, with no mention of patient characteristics. However, from the studies that are pertinent, few clear patterns emerge

concerning the relation between process measures of quality and such characteristics as age, race, sex, or income.

By age of patient. Findings on the relation between age of patients and process measures of quality are among the most contradictory. We have studies that show that older patients receive better care. Using five different measures of quality that included both process and outcome measures, Brook evaluated the care provided to patients with three different conditions: urinary tract infection, hypertension, and ulcers. Only one clear correlation between quality and demographic characteristics emerged, that between quality and the patient's age: "Irrespective of condition . . . there appears to be a definite significant relationship between age and all measures of quality of care (the older the patient, the better the quality of care)."[40] Similarly, a study conducted at five hospitals in Baltimore showed that older patients had a lower proportion of appendectomies classified as "unnecessary" or "doubtful" than did younger patients.[41]

However, in a study of the appropriateness of injections given by physicians in New Mexico, Brook reports results that are counter to the clear pattern that emerged from his earlier investigation: Brook and Williams[42] found that younger people, and especially children, were given fewer injections judged to be inappropriate than older persons. For patients under 14, pediatricians gave one-third as many injections judged inappropriate as either general practitioners or internists treating patients in that age group. But even among general practitioners who treated both children and adults the number of inappropriate injections given to children was about one-half that for adults.

Other reports of younger patients receiving more appropriate care include the study by Rubenstein et al. of women treated for symptoms of acute urinary tract infection at the emergency rooms of a voluntary hospital and a municipal hospital in the Bronx. At both hospitals process measures of quality differed significantly in relation to the three age groups examined (25 years or under, 26–34, 35 and over): the older the patient, the lower the quality of care.[43] Similarly, a study of appropriateness of drug prescriptions in a rural county showed that patients under 25 years of age not only received fewer prescriptions than other age groups, but that the percentage of drugs judged more appropriate was significantly higher for that age group than for all older age groups.[44]

In some studies both patterns were found to be present: for certain conditions younger patients received significantly higher quality care, but for other conditions older patients were the ones receiving the better care. A recent study of ambulatory care reports higher scores for older patients in two diagnostic categories and lower ones in the two other categories they

studied.[45] Lyons and Payne,[23] for their part, found in a study of inpatient care that process measures of quality were higher in two diagnostic categories and lower in one for patients between the ages of 14 and 64 compared with those 65 and older. In the 12 other diagnostic categories they examined, there were no significant differences between those two age groups. They similarly found no differences in the quality of ambulatory care provided to those two groups.[24]

In terms of establishing whether process inequities related to age exist, the studies just cited lead us to the unsatisfactory conclusion that in some instances no such inequities exist while in others they do; when they are present, sometimes they favor the young, sometimes the old. Obviously, unless we are able to predict when inequities related to age will be present and what the direction of such inequities is likely to be, such a conclusion has limited meaning and therefore little practical value. Unfortunately, the empirical findings exhibit no clear pattern that we are able to discern. Neither the authors of the studies nor anyone else, as far as we know, have suggested any theories or generalizations that would account for the different relations observed between the age of patients and the quality of care. In several of the studies cited, in fact, no attempt is made to account for the findings about that relation. However, several of the studies do offer explanations of the findings, which are now briefly discussed.

Rubenstein et al.[43] suggest that younger females may have received better care for urinary tract infections because of their similarity in age with the house staff at the two emergency rooms included in the study. Lyons and Payne, on the other hand, attribute the higher quality of care scores in the care given to patients with bronchitis who were over 65, compared with younger patients with the same diagnosis, to the frequent association in older age between bronchitis and other conditions such as heart disease, emphysema, and diabetes. In their view such a situation prompts practitioners to pay special attention to the management of the older patient's bronchitis. On the other hand, they speculate that patients over 65 with cerebrovascular accidents receive lower quality care than younger ones because cerebrovascular accidents "are frequently terminal events in the elderly, and little, if any, manipulation or therapeutic intervention will alter the course of events . . . especially in a patient with multiple additional diagnoses or disabilities for whom survival may mean only greater disability and societal disruption."[23] The explanations given by Lyons and Payne suggest, in essence, that the care provided was in fact appropriate. From what they say, it can be concluded that the relatively few differences by age that they found were not isolated instances of inequity. Rather, they were an indication that the criteria they used were not sufficiently discriminating to cover the kind of special situations they describe.

Brook and Williams,[42] for their part, attribute the greater number of inappropriate injections given to those over the age of 45—injections that included a large proportion of nonspecific ("placebo-type") medications—partly to the expectation on the part of such patients that an injection will be given to them by the physician. The suggestion, then, is that physicians give in to the patient's expectations of what constitutes appropriate care. If so, that may also explain why general practitioners were found to give fewer inappropriate injections to children than to adults. Unlike adults who may expect to be given an injection, children are likely to express loud displeasure at the prospect of an injection, so that the physician gives shots to children only when it seems absolutely necessary. We shall return later to this and other explanations of differences in quality of care related to the patient's age.

By race of patient. With respect to differences in quality of care associated with the patients' race, the studies cited above, to the extent that they reported on it, showed only very few, isolated instances in which significant differences were found. One study that did yield clear differences by race is Janzen's evaluation of the quality of care provided by nurses in 19 Michigan hospitals.[46] The interaction between nurses and hospitalized patients was scored by a trained observer using a rating instrument that, because of the importance it accords to the "psychosocial" aspects of care, provides at least as much an assessment of the interpersonal aspects of care as of technical ones.

There were significant differences in the care provided to white patients as compared with "others" in all categories of the rating scale. To the extent that those results reflect the quality of interpersonal care, they provide another instance of a pattern we noted in our discussion of structural inequities that are related to race: although there is no clear evidence of inequities related to race when examining the technical aspects of care, it is relatively easy to find evidence for such inequities in terms of the interpersonal aspects of care, even though, paradoxically, the interpersonal aspects are the ones generally considered the less likely to be documented and the more refractory to evaluation.

Further instances of the pattern just described are found in evaluations of nursing home care that concentrate on interpersonal aspects. In one such evaluation of care in a nursing home, Gottesman and Bourestom[47] found that race was one of the factors correlated with the overall quality of care, which they measured in terms of the amount of the patient's interaction with the staff and with other residents. By that interpersonal care-oriented measure, nursing homes with a preponderance of white patients were found to provide higher quality care.

By sex of patient. As with race, few significant differences according to sex of the patient have been reported. Hulka and associates[45] found, in terms of their process measure of quality, that although there were no differences by sex with respect to the care given to patients with diabetes and hypertension, medical examinations given to men received significantly higher scores than those given to women. Although the authors do not comment on this finding, Mudd and Fleiss[48] report in a study of physical examinations given to hospitalized adults that performance on those examinations differed by the sex of the patient only with respect to the completion of genital and rectal examinations. The authors attribute the less frequent performance of rectal and genital examinations on women to a number of reasons, including lack of time and embarrassment on the part of the physician. Because all physicians included in the study by Hulka and associates were men it is possible that the results are at least partly the result of physician inhibitions.

Some of the differences in quality of care that have been reported in relation to sex actually combine sex and age. For example, Brook and Williams found that women, after the age of 34, received more injections judged inappropriate than men of the same age. The authors attribute that difference to the inappropriate use of injectable estrogens and progesterones to treat "nonspecific symptoms associated either with female aging or with the menopausal syndrome."[42]

Janzen,[48] in the previously mentioned study of the quality of care provided by nurses, found that in general the quality of nursing care was better for younger males than older ones, whereas it was better for older women than younger ones. Although no significant differences by sex or age alone were found, the interaction effect of age and sex was found to be significant for one-half of the components of the quality of care measure used. No explanations of this pattern are offered; it would be useful to know whether these results are related to the characteristics of the providers of care, in this case nurses who are likely to be young women.

Thus, with respect to quality of care inequities related to the patient's sex, the most that can be said, based on the fragmentary evidence available, is that although there is no evidence of pervasive inequities by sex, instances of inequity do exist. The traditional male predominance in medicine and female predominance in nursing have often been considered a source of inequitable treatment. The literature that began appearing in the early 1970s as part of the women's health movement attests to a profound dissatisfaction directed primarily, though not exclusively, at the interpersonal aspects of medical care provided to women.[49,50] We have cited a few examples in which the match between the provider's age or sex and that of

the patient might account for differences in both the technical and interpersonal aspects of care. However, in the absence of more systematic studies of this relation, no firm conclusions can be drawn about the inequities that may be associated with it.

By patient's income. Finally, when we turn to the evidence on process inequities related to income and ability to pay, we again find that the patient's income or ability to pay for care was not included in some of the more extensive studies assessing the process of care, but those studies in which income or ability to pay were included report no significant correlations between those factors and quality of care. Brook[40] found that Medicaid status did not correlate with process measures of quality of care, and Payne and associates reported that, with respect to care provided to hospitalized patients in Hawaii, "expected source of payment is not a consistent or predictable correlate of physician performance."[4]

Analogous findings were reported from a set of relatively large-scale studies of Office of Economic Opportunity (OEO) neighborhood health centers.[51-53] In our discussion of structural inequities we dealt with some of the sources of care most commonly used by lower socioeconomic groups, including, in particular, outpatient departments, city clinics, and private practices with a large proportion of Medicaid patients. Another category of such facilities is what has come to be known as the community health center. These are government sponsored and are intended to improve access to health services for underserved populations, particularly in the inner cities and rural areas.

OEO neighborhood health centers were the first type of community health center to be established. They were evaluated, using process measures of care, by Morehead and her associates,[51-53] who reported that, on the whole, the quality of the care provided at the neighborhood health centers they studied was similar to that provided in the outpatient departments of hospitals affiliated with medical schools and in private group practices. The relative performance in those three settings were judged to be about equal, an indication that the low-income population served by the centers was not, at least in comparison with those two other settings, receiving inequitable treatment with respect to quality of care. On an absolute scale, however, performance was not very high in any of the settings studied, even by the relatively lenient standards used by Morehead and associates. In the intervening decade since those studies were completed, neighborhood health centers have undergone many changes, including, in particular, a movement toward serving a wider cross section of the population; little is known about the effect of those changes on the quality of care provided by those centers to low-income persons.

The studies cited so far, then, provide some evidence that, perhaps contrary to commonly held assumptions, lower-income groups and those less able to pay for their care do not receive lower quality care. There is evidence, in fact, that those who are better able to pay for their care may at times receive worse care, usually in the form of unnecessary procedures. Thus Sparling[41] found that, in the three community hospitals he studied, the percentage of operations judged to be either "unnecessary" or "doubtful" varied in accordance to the extent and type of coverage for hospital expenses. The percentage of unnecessary or doubtful operations was lowest for patients on welfare, followed by private pay patients, patients with insurance other than Blue Cross, and finally those covered by Blue Cross. Kessner et al.[54] found a strong inverse relationship between tonsillectomies and six measures of social class, including, in particular, family income. Although they did not actually evaluate, as Sparling did, the necessity for the operations, they note that the large variation in tonsillectomy rates they found "is disturbing because clinical indications for tonsillectomy, in general, are limited as are differences in indications among children . . . Because the overall tonsillectomy rates are related to family income, the data suggest that ability to pay, rather than medical need, may well be an important deciding factor for whether the child will keep or lose his tonsils."[54]

Process inequities: summary and discussion. Given the many contradictions and unexplained relationships in the evidence we have just reviewed, it is difficult to draw any firm conclusions concerning the existence and the nature of process inequities. We are even reluctant to conclude, in spite of the findings that show few significant differences in quality by patient's age, sex, race, or income, that indeed there are no inequities related to those factors. It is all too likely that the criteria lists used in most assessments of the process of care are often not sufficiently refined to enable us to detect certain types of inequities. By the same token, we cited one example in which Lyons and Payne[23] could account for differences in quality of care by age in a way that made those differences seem appropriate; therefore, where inequity might have been suspected, there was none.

Similarly, when Hulka and associates[45] compared process measures of quality that they obtained using a criteria list with the judgments of expert physicians, they found cases in which the care provided had received a low score based on the criteria, but high scores were given by the judges because the latter were able to take into consideration certain special characteristics of the patients. Such discrepancies are particularly important from the perspective of equity when they go in the other direction: when certain patients—the poor, the elderly, the uneducated—have special needs that are *not* reflected in the criteria lists, so that their care is apt to receive scores equal to that of

other groups, and yet be deficient in relation to their special needs.

In addition, the fragmentary evidence we have does show that some process inequities exist with respect to the interpersonal aspects of care. And finally, it should, perhaps, be remembered that inequities sometimes affect the privileged, in a way that exposes them to potential harm while saddling society with often wasteful use of valuable resources. For all those reasons, we believe it would not be justified to conclude, even if the available data point in that direction, that process inequities are not commonly found.

Outcome Inequities. Outcome inequities are the third and fourth form of inequity identified in our framework. By our definition, outcome inequities are present whenever any one group, compared with other groups, experiences a larger proportion of unfavorable health outcomes that can be attributed to care.

It may appear at first that with respect to health outcomes the data documenting inequities are plentiful and unambiguous. For example, the data on the health status of low-income groups assembled by Rudov and Santangelo[55] provide ample documentation of the disparities that exist with respect to numerous measures of health when minorities and low-income groups are compared with the rest of the population. However, those disparities (such as differences in immunization rates) in most instances represent shortcomings not in quality of care per se, but in the quantity of care, and therefore in basic access to care. In addition, many differences in mortality rates and similar indicators are the result of a number of factors, of which health care is only one, and often a relatively minor one at that. Because our concern, as we stated at the outset, is more narrowly confined to quality of care, our interest focuses only on those differences in health outcomes that can be attributed to care received. As with the other aspects of quality that are of concern to us here, the data that relate to such differences are not easily found, making it difficult to determine the nature and extent of outcome inequities present.

A major study of surgical outcomes was conducted a few years ago by the staff of the Stanford Center for Health Care Research. In that study the quality of surgical care was measured in terms of postsurgical mortality and morbidity, adjusted for differences among patients in presurgical condition. Based on the data from that study, Flood and Scott[56] report that only income was significantly related to patient outcomes: patients with higher incomes tended to have better surgical outcomes. No relation was found between outcomes and any of the other patient characteristics they measured, which included ethnicity, marital status, education, and social stress.

Thus, with respect to surgery at least, higher income groups simultaneously benefit and suffer from inequities in quality of care. As we saw earlier, they are subject to more unnecessary surgery, but apparently surgery performed on them, whether it is necessary or not, is done with better results than for other groups.

By contrast, when Kessner et al.[54] evaluated whether children who needed glasses had them, and, of those children who had them, whether the glasses were needed and were the correct prescription, they found equally poor performance at all socioeconomic levels. They conclude that "it is clear that a high socioeconomic status does not guarantee effective management of visual disorders." We thus have here a stark example of equity, where everyone received equally poor care.

As mentioned earlier, differences in *preventable* maternal and infant mortality rates represent a measure of long standing that suits our need well for measures that reflect outcomes attributable to care. In addition, such measures are also related to the mortality indices so often cited as proof of the disparities that continue to exist between different races and income groups. Unfortunately, we were not able to identify any recent reports of data on preventable mortality rates. Schaffner and associates,[57] in their report on direct maternal mortality in Michigan between 1950 and 1971, show that the disparities between whites and nonwhites did not change over the 22-year period. They do not provide, however, a breakdown by race of their data on percentage of direct maternal deaths classified as preventable; therefore we cannot judge how great the inequities by race are in relation to quality of care, nor whether they have changed over time.

By the same token, we have no basis for evaluating whether Schaffner and associates are correct when they state that "it could be inferred that the prenatal obstetrical care available to the nonwhite pregnant woman lags over 22 years behind that available to her white counterpart." Although a reasonable inference, it is not consistent with the relationships observed by Hooker[58] in his pioneering study of maternal mortality in New York City. Hooker found that the maternal death rate among black women was twice that for white women, but the differences in preventability of deaths between the two groups were very small: 67.7% for black women, and 65.5% for white women.

As with the other types of inequity, then, we have evidence that some outcome inequities exist, but even the limited evidence we identified has gaps and contradictions that make any generalizations difficult. However, an important point we have tried to illustrate, and one to which we shall return in the concluding section, is that differences in health status or in

mortality rates are not necessarily evidence that inequities in the *quality of care* provided are also present.

Summary and Discussion

In this final section we briefly summarize what the evidence we reviewed suggests in terms of the framework we developed for identifying inequities in the distribution of quality of care. We then discuss the need for additional research and data, as well as the policy issues related to this area of concern.

Summary and Discussion of the Evidence. As should be clear by now from our examination of the evidence that relates to the question of equity in the distribution of quality of care, it is not possible for us to proceed through our framework for examining this question shown in Figure 2 and identify the specific structural, process, and outcome inequities that exist, specifying for each type of inequity what groups are affected and in what way. There are nevertheless some inferences that can be drawn from the evidence we reviewed.

One such inference is that inequities, when present, do not necessarily work to the detriment of such traditionally disadvantaged groups as the poor. In that respect we agree with Brook and Williams:

> The evidence does not support the hypothesis that the technical quality of health care provided to the disadvantaged is uniformly or universally poorer than that provided for other persons. In some settings, or for certain groups within the disadvantaged population, quality may be even better.[8]

However, although *technical* quality may not always be worse for such disadvantaged groups, the available evidence suggests that for them the *interpersonal* aspect of quality usually is. The traditional ward accommodations and the outpatient facilities of large teaching hospitals are one example of structural inequities characterized by technical quality that may be better, but are accompanied by quality that may be worse from the standpoint of interpersonal relations.

Similarly, we noted that better access to care for the more privileged may expose them to care that is worse because it is unnecessary. The privileged also are not always protected from inferior care, as the example of the management of visual disorders in children demonstrates.[34]

Another point that warrants repetition is that structural and other *differences* among groups are not necessarily associated with inequities related to *quality of care*. Thus, although different income groups may receive care from different sources, the care they receive often appears to be equally appropriate, at least to the extent that differences are detectable with the measures of technical quality used. As we have indicated,

however, the evidence provides less support for the same argument with respect to the interpersonal component of care.

Because ideally there should be no inequities with respect to either the technical or the interpersonal aspects of quality—especially if those are seen as two mutually reinforcing elements—it may well be that, from the point of view of structural arrangements, there are few instances in which it could be claimed that clear structural differences do not translate into structural inequities, at least with respect to the interpersonal component of care. With respect to quality of care, as in other areas of life, it may be difficult to marshal strong empirical evidence in support of a "separate but equal" argument.

For another aspect of differences among groups, we have argued that differences among groups with respect to health status or to outcomes such as mortality rates do not necessarily reflect inequities in terms of the quality or even the quantity of care provided. We believe it to be an important and valid point, even though we could cite only one study, and one done very long ago at that, that clearly supports that argument in terms of quality of care. However, it is a question that is also related to the broader argument about the contribution of health-care services to health, and we return to that argument in the final subsection of this paper.

Finally, in the findings presented there is some evidence that traditional measures of quality of care, and particularly process measures based on criteria lists, useful though they are in many respects, may not be sensitive enough to detect the kinds of inequities that are of interest here. On one hand, the use of such lists may point to inequities in quality that, on closer scrutiny, reflect a course of action that appropriately takes into account the particular characteristics and needs of the patient. On the other hand, we can easily conceive of results that show no inequities when they in fact exist but were not detected because the evaluation criteria did not include the explicit consideration of the special clinical and other needs engendered by the age, income, education, and socioeconomic status of the patient. Thus at least some of the contradictions and unexplained patterns in the findings we reviewed may be attributed to the deficiencies that certain traditional quality of care measures have when they are pressed into service for the identification of inequities in the distribution of quality of care.

Research and Data Needs. The question of whether there are inequities in the distribution of quality of care and what those inequities imply is complex and not subject to simple answers. It is, as we have noted repeatedly, all the more difficult to address that question given the paucity of directly relevant studies. However, by outlining in this section the kinds of studies we believe should be conducted, we do not mean to imply that those studies of themselves will provide all the

answers that currently elude us. More likely than not they will raise new, additional questions. However, in the absence of the necessary evidence, no answers to any questions can be expected in the near future.

As we have noted, the study of the related area of inequities in the distribution of the *quantity* of health-care services has been greatly facilitated by the major data collection and data analysis efforts that the National Health Interview Survey and the CHAS/NORC surveys represent. Those sources have in fact provided the data that point to an apparent reduction in the inequity of access to services, which in turn provides the basis for considering the next step, that of examining how equitably the *quality* of those services is distributed—the concern of this paper. However, in order to truly move to that next step, data are needed that are parallel in breadth and scope to those that are now available on the quantity of services used.

More specifically, data are needed that allow meaningful quality of care comparisons across different regions and across different population groups. Virtually all the studies on quality of care currently available focus on a single institution or a single city or county, making it difficult to generalize. Furthermore, if the interest is in issues of equity, studies of quality of care that are narrowly focused on disadvantaged populations, such as those cited by Brook and Williams,[8] are of limited value, because differences in assessment techniques, if nothing else, make it difficult to compare the results of studies of the disadvantaged with studies of nondisadvantaged populations.

In addition, studies that reflect a recognition of the relations between structure, process, and outcome are apt to provide greater insight into the nature of any quality of care inequities that may be present. Evaluations of quality in terms of any of those three elements—structure, process, or outcome—can only gain from the use of evaluation criteria that reflect what is known, and not known, about the relations among the three elements. The evaluation criteria, as we have mentioned before, should also reflect, to the extent possible, special considerations in relation to such characteristics of the patient as education, income, and area of residence. The resulting evaluations are more likely to reflect a conception of quality that includes the consideration of the broadest possible spectrum of the patient's needs.

We have also noted that the patient's view of what is appropriate care may come into conflict with that of the practitioner and result in services that could be classed as inappropriate as well as dangerous and wasteful. We need more information about this topic to help us come to grips with the equity considerations that relate to it. It is possible that in the near future some of the necessary information and analytic tools on this issue will emerge from the rapidly growing literature on clini-

cal decisionmaking, some of which includes the consideration of patients' and physicians' utilities and preferences.

We have also noted that the interpersonal aspect of care is especially important in any consideration of equity because it seems to be more refractory to change and modification. The CHAS/NORC studies already include many questions designed to measure patient satisfaction, although the focus, as is appropriate in studies of access, has been more on the amenities of care than on the care itself. A greater emphasis in those surveys on the interpersonal aspects of care, and on the patient's view of technical quality, would yield very valuable information in the consideration of equity in the distribution of quality of care.

A question related to the interpersonal aspect of care that may warrant special attention, based on some of the studies we reviewed, is the question of how important for quality of care is the congruence between the practitioner's and the patient's characteristics, particularly with respect to age, sex, and race. It is a topic that is often mentioned, and sometimes hotly debated, but it has not received the comprehensive study and attention it deserves.

Other areas that deserve special attention in the study of quality of care from the perspective of equity are those that are farther away from the well-worn path of acute care services, and in which some of the more glaring inequities in terms of both the amount and type of services appear to exist: dental care, nursing-home care, and mental health services.

Needless to say, our preference is for a continuing series of studies, much like the CHAS/NORC series, to measure changes over time, while facilitating successive refinements in the methodologies used to obtain the data. Special, more comprehensive studies could also be conducted better in the context of such a basic series of studies directed at quality of care. Such special studies might address the central question of what the factors are that account for any inequities that are found to exist; answers to that question are essential to any systematic development of corrective policies.

It may be asked whether the kind of large-scale data collection effort we propose is justifiable, especially at a time when the resources to mount such an effort may be particularly scarce. It may not be. But until an effort of that magnitude is undertaken, we continue to know very little about how the quality of care received by various disadvantaged groups compares with that received by the general public. That is not a question of merely academic interest, because to assure access to health services for persons below the poverty line the Federal government has spent over $20 billion this past year.[59]

Policy Implications. Although the currently available data do not make it possible to identify the exact extent and nature of

inequities in the distribution of quality of care in the United States, there is sufficient evidence that inequities do exist to raise the question, for policy makers, of what should be done about it. The question may become especially pressing if more becomes known about such inequities as a result of the kind of research and data collection effort we described in the preceding subsection.

However, from a policy perspective, a prior question may be whether any effort at all should be directed at eliminating inequities in the distribution of quality of care. A position against any intervention can be taken on purely ideologic grounds, reflecting the view that the removal of those inequities should not be achieved through concerted, collective action, and especially not through government intervention. There are, however, other arguments that can be made for not eliminating quality of care inequities. The two such arguments we discuss are:

(1) Given the overall low level of quality of care, the concern with inequities is misplaced and all efforts should be directed at the far more important and pressing task of raising quality of care overall;

(2) concerns with inequities in quality of care assume that health status can be substantially affected by the quality of health services; that, however, is an unproven and questionable assumption, especially for the very people who are the subject of inequities, such as the poor and the aged.

Brook and Williams stress not the relative but the absolute level of quality of care provided to disadvantaged populations when they state that "the quality of technical health care provided for the American people in general suffers from a number of deficiencies . . . To say that the quality of care for the disadvantaged compares reasonably well with a less-than-optimal product is not a particularly positive statement."[8] One might conclude, therefore—although Brook and Williams do not go that far—that first priority should be given to improving overall care because the relative differences in quality pale by comparison with the low absolute level of care currently being provided.

The overall level of quality of care, as measured in all the studies we cited, is indeed often lower than we would desire. Serious and less serious deficiencies are repeatedly noted in process assessments of quality of care. It is also true, as we have seen, that some structural inequities actually benefit those who use the less desirable sources of care, and that programs specifically set up to serve the disadvantaged, such as OEO neighborhood health centers, do not appear to provide care that is much different in its quality from that provided in other settings.

We have also seen, however, that Medicaid patients are served not by a cross section of physicians but by a specific segment of generally less qualified practitioners. To raise the level of such practitioners' care is likely to require more effort and creativity than it would for other physicians. We have also suggested that if quality is measured in terms of its technical *and* interpersonal aspects (as we believe it should be) the goal of "separate but equal" facilities with respect to quality of care is, from a practical point of view, not very easily attained. In that sense we see support in the studies we reviewed for a principle that Dutton derived from her study of ambulatory care services in Washington, D.C.:

> If diverse health care systems continue to be available which, based on the net effect of their particular advantages and disadvantages, are inherently more or less desirable, the more desirable systems will be used by the more privileged members of society, while the less desirable systems are allocated to the less privileged.[30]

It may well be, then, that the choice is not between working on eradicating inequities in the quality of care or improving quality of care in general. We see continuing evidence that both need to be done, and the challenge lies in devising a strategy that will make it possible to progress simultaneously on both fronts.

There is, however, an even more fundamental basis for challenging any effort to reduce or eliminate inequities in the distribution of quality of care. It can be claimed that we have reached a point in our society where the marginal gains to be derived from more or better care are very small, but the costs of such improvements are very high. As for the disadvantaged, even if they receive a lower quality of care than the rest of the population, it can be argued that health care is not the critical factor that accounts for their demonstrably lower health status. Effectively reducing the much higher infant mortality rates of certain ethnic and minority groups, it has been argued, will take more than building another sophisticated neonatal unit or even providing better prenatal care.[60,61]

The argument that Starr has termed "therapeutic nihilism,"[62] which holds that health care is on balance more harmful than beneficial, is not widely shared, but there is growing recognition that much of what is done by physicians and other practitioners is not positively known to affect the course of disease.[63]

In fact this uncertainty about the relative efficacy of various treatments hampers the evaluation of quality of care to the extent that it forces an incongruous reliance on process measures even when those are found repeatedly not highly related to outcomes. In addition, as we have noted, outcomes are influ-

enced by a list of factors of which the quality and quantity of care are only one set.

To conclude, however, that quality of care should not be a concern at all could be going too far. A more reasonable course of action may be to direct efforts designed to reduce inequities in quality primarily at those areas where there is the greatest certainty that care does affect outcomes significantly. In addition, as Davis and Schoen[7] have suggested, there must be more knowledge and understanding of the impact of various programs directed at improving the health status of certain population groups. In general, there must be more known about the nonmedical factors most likely to potentiate the effects of the care provided to groups that have the lowest health status. Knowledge of those factors would enable providers to work in concert with other professionals and agencies to raise the health status of our most deprived citizens. Such a view of the practitioner's role only reflects, after all, one of the elements of good medical care that Lee and Jones[2] identified long ago.

Finally, if the intent is indeed to eliminate distributional inequities with respect to quality of care, what guidance does our review provide? As we see it, the framework in terms of which we structured this review (Figure 2) is a useful guide for devising a general strategy to reduce these inequities.

Thus, before being concerned with inequities related to quality of care, it must be ascertained that there is effective access to the amount of care appropriate to a patient's needs. The next step, consistent with our framework and our earlier discussion, is to eliminate structural inequities, particularly those that result in clearly lower levels of quality for any group, where quality is measured both in terms of the technical and interpersonal aspects of care.

Once equity with respect to the appropriateness of the providers has been achieved, the next step is to remedy any process inequities that persist. Two areas are of special concern here. One is unequal treatment of different groups by the same provider, which reflects discriminatory behavior. Remedying such inequities may prove a very difficult challenge, especially with respect to achieving equity in the interpersonal aspects of care. Even where there is no apparent discrimination, however, as in some prepaid group practices, several studies have shown that certain groups—such as blacks and lower income groups—tend to use services differently from others, such as by their tendency to use drop-in clinics and emergency rooms rather than make appointments.[30,64] To what extent such differences represent inequities and should therefore be actively eliminated must depend, consistent with what we have said before, primarily on how those differences affect the technical and interpersonal aspects of the quality of care received.

The other aspect of process inequities that is of special concern is the need to ascertain that any process measures of quality used in monitoring process inequities take into account, to the extent possible, the special clinical needs of those most likely to suffer from inequities. As part of that, outcome inequities must be examined for evidence that process measures are not adequately accounting for such special needs.

We can thus outline a general strategy, but it is difficult to suggest the specific mechanisms that will bring about all the changes we have outlined. It is difficult to disagree with Brook and Williams[8] when they conclude that efforts to improve the care provided to the disadvantaged have not produced satisfactory results. Many of the types of inequities that we have described have roots that extend deep into the structure of our society, seemingly out of reach of those whose grasp is only within the confines of medical care. The challenge is a difficult one and may require some experimentation with fresh approaches, as Brook and Williams suggest. It must also involve the public sector, if only because vast sums are being spent by state and Federal governments to assure access to care for the poor, the elderly, and other special groups. Such public financing to provide access to care carries with it the responsibility of at least ensuring that the care is safe; there is the additional responsibility to ensure that the services being paid for are of acceptable quality. The creation by Congress of Professional Standards Review Organizations (PSROs) represents a recognition by the public sector of the responsibility for the quality of care purchased with public funds. Although PSROs have had an effect on the excessive use of services, a modest positive effect on quality of care is only beginning to become apparent.[65]

Finally, we emphasize once again that in our view efforts undertaken from any quarter to reduce inequities in the distribution of quality of care must be based on a better understanding of the factors at work. Even as efforts proceed to reduce remaining inequalities with respect to access—following the categorical program approach suggested by Aday et al.[3] or another approach—and as the focus of those efforts then shifts to quality, it is imperative that there be a corresponding effort in data collection and research that addresses the question of equity in the distribution of quality of care.

References

[1] United States Special Committee on Aging, Subcommittee on Long Term Care. FRAUD AND ABUSE AMONG PRACTITIONERS PARTICIPATING IN THE MEDICAID PROGRAM. Washington, DC: Government Printing Office, 1976.

[2] Lee RI, Jones LW. THE FUNDAMENTALS OF GOOD MEDICAL CARE. Publications of the Committee on the Costs of Medical Care (no. 22). Chicago: University of Chicago Press, 1933.

[3] Aday, LA, Andersen R, Fleming GV. HEALTH CARE IN THE U.S.: EQUITABLE FOR WHOM? Beverly Hills, CA: Sage Publications, 1980.

[4] Payne BC, Lyons TF, Dwarshius L, Kolton M, Morris W. THE QUALITY OF MEDICAL CARE: EVALUATION AND IMPROVEMENT. Chicago: Hospital Research and Educational Trust, 1976.

[5] Hart JT. *The inverse care law.* LANCET 1971;1:405.

[6] Andersen R, Kravitz J, Anderson OW, eds. EQUITY IN HEALTH SERVICES: EMPIRICAL ANALYSES IN SOCIAL POLICY. Cambridge, MA: Ballinger, 1975.

[7] Davis K, Schoen C. HEALTH AND THE WAR ON POVERTY: A TEN-YEAR APPRAISAL. Washington, DC: The Brookings Institution, 1978.

[8] Brook RH, Williams KN. *Quality of health care for the disadvantaged.* J COMMUNITY HEALTH 1975;1:132.

[9] Verbrugge LM. *Equity of access to health care: a research review and prospectus.* In: U.S. HEALTH GOALS FOR THE YEAR 2000, background papers. Washington, DC: Institute of Medicine, National Academy of Sciences (in press).

[10] Donabedian A. *Effects of Medicare and Medicaid on access to and quality of health care.* PUBLIC HEALTH REP 1976;91:322.

[11] Donabedian A. *Evaluating the quality of medical care.* MILBANK MEM FUND Q 1966;44:166.

[12] Donabedian A. A GUIDE TO MEDICAL CARE ADMINISTRATION, VOLUME II: MEDICAL CARE APPRAISAL—QUALITY AND UTILIZATION. New York: American Public Health Association, 1969.

[13] Donabedian A. *Needed research in the assessment and monitoring of the quality of medical care.* Hyattsville, MD: U.S. Department of Health, Education, and Welfare (Publication No. (PHS) 78-3219), 1978.

[14] Donabedian A. THE DEFINITION OF QUALITY AND APPROACHES TO ITS ASSESSMENT. Ann Arbor, MI: Health Administration Press, 1980.

[15] Brook RH. *Critical issues in the assessment of quality of care and their relationship to HMOs.* J MED EDUC 1973;48:114.

[16] Donabedian A. *A frame of reference.* QRB 1976;2:5.

[17] Donabedian A. *Measuring and evaluating hospital and medical care.* BULL NY ACAD MED 1976;52:51.

[18] Brook RH, Williams KH, Avery AD. *Quality assurance in the 20th century: will it lead to improved health in the 21st?* In: Egdahl RH, Gertman PM, eds. QUALITY ASSURANCE IN HEALTH CARE. Germantown, MD: Aspen Systems Corporation, 1976.

[19] Donabedian A. CONCEPTS OF HEALTH CARE QUALITY: A PERSPECTIVE. Washington, DC: Institute of Medicine, National Academy of Sciences, 1974.

[20] Riedel, RL, Riedel DC. PRACTICE AND PERFORMANCE: AN ASSESSMENT OF AMBULATORY CARE. Ann Arbor, MI: Health Administration Press, 1979.

[21] Payne BC, Lyons TF, Neuhaus E, Kolton M, Dwarshius L. METHOD OF EVALUATING AND IMPROVING AMBULATORY MEDICAL CARE. Final report. Ann Arbor, MI: The University of Michigan, 1978.

[22] Becker MH, Haefner DP, Kasl SV, Kirscht JP, Maiman LA, Rosenstock IM. *Selected psycHosocial models and correlates of individual health-related behaviors.* MED CARE 1977;15:27.

[23] Lyons TF, Payne BC. *The quality of physicians' health care performance—a comparison against optimal criteria for treatment of the elderly and younger adults in community hospitals.* JAMA 1974;227:925.

[24] Lyons TF, Payne BC. *Quality of physicians' office-care performance—different for elderly than younger adults?* JAMA 1974;229:1621.

[25] Weissert WG. *Long-term care: an overview.* In: Health United States 1978. Hyattsville, MD: U.S. Department of Health, Education, and Welfare (Publication No. (PHS) 78-1232), 1978.

[26] Egbert LD, Rothman IL. *Relations between the race and economic status of patients and who performs their surgery.* N Engl J Med 1977;297:90.

[27] Perkoff GT, Anderson M. *Relationship between demographic characteristics, patient's chief complaint, and medical care destination in an emergency room.* Med Care 1970;8:309.

[28] Kohl SG. Perinatal mortality in New York City: responsible factors. Cambridge, MA: Harvard University Press, 1955.

[29] Duff RS, Hollingshead AB. Sickness and society. New York: Harper and Row, 1968.

[30] Dutton DB. *Patterns of ambulatory health care in five different delivery systems.* Med Care 1979;17:221.

[31] Bellin LE. *Medicaid in New York: utopianism and bare knuckles in public health. Realpolitik in the health care arena: standard setting of professional services.* Am J Public Health 1969;59:820.

[32] Kavaler F. *Medicaid in New York: utopianism and bare knuckles in public health. People, providers and payment—telling it how it is.* Am J Public Health 1969;59:825.

[33] Bellin LE, Kavaler F. *Policing publicly funded health care for poor quality, overutilization and fraud—the New York City Medicaid experience.* Am J Public Health 1970;60:811.

[34] Bellin LE, Kavaler F. *Medicaid practitioner abuses and excuses vs. counterstrategy of the New York City Health Department.* Am J Public Health 1971;61:2201.

[35] Rosenberg SN, Gunston C, Berenson L, Klein A. *An eclectic approach to quality control in fee-for-service health care: the New York City Medicaid experience.* Am J Public Health 1976;66:21.

[36] Mitchell JB, Cromwell J. *Medicaid mills: fact or fiction.* Health Care Financing Review 1980;2:37.

[37] United States General Accounting Office. *A federal strategy is needed to help improve medical and dental care in prisons and jails.* Washington, DC: General Accounting Office, 1978.

[38] Myers JK, Schaffer L. *Social stratification and psychiatric practice: a study of an outpatient clinic.* Am Sociol Rev 1954;19:307.

[39] Hollingshead AB, Redlich FC. Social class and mental illness: a community study. New York: John Wiley and Sons, 1978.

[40] Brook RH. *Quality of care assessment: a comparison of five methods of peer review.* Washington, DC: US Department of Health, Education, and Welfare, 1973.

[41] Sparling F. *Measuring medical care quality: a comparative study.* Hospitals 1962;36:62.

[42] Brook RH, Williams KN. *Evaluation of the New Mexico peer review systems, 1971 to 1973.* Med Care 1976;14 (Supplement to No. 12).

[43] Rubenstein L, Mates S, Sidel VW. *Quality-of-care assessment by process and outcome scoring—use of weighted algorithmic assessment criteria for evaluation of emergency room care of women with symptoms of urinary tract infection.* Ann Intern Med 1977;86:617.

[44] Miles DL. *Multiple prescriptions and drug appropriateness.* Health Serv Res 1977;12:3.

[45] Hulka B, Romm FJ, Parkerson GR, Russell IT, Clapp NE, Johnson FS. *Peer review in ambulatory care: use of explicit criteria and implicit judgments.* Med Care 1979;17 (Supplement to No. 3).

[46] Janzen E. *Quality nursing care assurance.* Unpublished manuscript, Wayne State University, 1973.

[47] Gottesman LE, Bourestom NC. *Why nursing homes do what they do.* Gerontologist 1974;14:501.

[48] Mudd JW, Fleiss JL. *Physical examinations of hospitalized adults.* J Med Educ 1973;48:1140.

[49] Franfort E. Vaginal politics. New York: Bantam, 1973.

[50] Sprague JB. *Women and health bookshelf.* Am J Public Health 1975;65:741.

[51] Morehead MA. *Evaluating quality of medical care in the neighborhood health center program of the Office of Economic Opportunity.* Med Care 1970;8:118.

[52] Morehead MA, Donaldson RS, Sevaralli M. *Comparisons between OEO neighborhood health centers and other health care providers of ratings of the quality of health care.* Am J Public Health 1971;61:1294.

[53] Morehead MA, Donaldson RS. *Quality of clinical management of disease in comprehensive neighborhood health centers.* Med Care 1974;12:301.

[54] Kessner DM, Snow CK, Singer J. Assessment of medical care for children, Washington, DC: Institute of Medicine, National Academy of Sciences, 1974.

[55] Rudov MH, Santangelo N. *Health status of minorities and low-income groups.* Washington, DC: US Department of Health, Education, and Welfare (Publication No. (HRA) 79-627), 1979.

[56] Flood AB, Scott WR. Professional power and professional effectiveness: the power of the surgical staff and the quality of surgical care in hospitals. Stanford, CA: Stanford University, 1977.

[57] Schaffner W, Federspiel CF, Fulton ML, Gilbert DG, Stevenson LB. *Maternal mortality in Michigan: an epidemiologic analysis, 1950–1971.* Am J Public Health 1977;67:821.

[58] Hooker RS. Maternal mortality in New York City, a study of all puerperal deaths, 1930–1932. New York: The Commonwealth Fund, 1933.

[59] Executive Office of the President, Office of Management and Budget. The budget of the United States Government, Fiscal Year 1981. Washington, DC: Government Printing Office, 1980.

[60] Weaver JL. *Policy responses to complex issues—the case of black infant mortality.* J HEALTH POLIT POLICY LAW 1977;1:433.

[61] Boone MS. *Fertility, disease, and infant mortality in an American inner city population.* Paper presented to the Society for Applied Anthropology, Denver, CO. Washington, DC: Georgetown University, 1980.

[62] Starr P. *The politics of therapeutic nihilism.* WORKING PAPERS FOR A NEW SOCIETY 1976;4:48.

[63] Cochrane AL. EFFECTIVENESS AND EFFICIENCY: RANDOM REFLECTIONS ON HEALTH SERVICES. London: The Nuffield Provincial Hospitals Trust, 1972.

[64] Nolan RL, Schwartz GL, Simonian K. *Social class differences in utilization of pediatric services in a prepaid direct service medical care program.* AM J PUBLIC HEALTH 1967;57:34.

[65] Health Care Financing Administration, *Professional Standards Review Organization—1979 program evaulation.* Washington, DC: Health Care Financing Administration, 1980.

[66] Donabedian A, Axelrod SJ, Wyszewianski L, eds. MEDICAL CARE CHARTBOOK. 7th ed. Ann Arbor, MI: AUPHA Press, 1980.

Health Status as a Measure of Need for Medical Care: A Critique

John Yergan, M.D.*
James LoGerfo, M.D., M.P.H.*†
Stephen Shortell, Ph.D.†
Marilyn Bergner, Ph.D.†
Paula Diehr, Ph.D.‡
William Richardson, Ph.D.†

Since the early part of this century, there has been considerable interest in relating access to medical care to the health needs of the population. Underlying this interest has been the desire to assure that people have ease of access to personal medical care services that will maintain health, restore it, or at least slow its deterioration. Despite the high level of interest, however, there has been considerable debate over how to measure either access to care or the health status of an individual. Given the lack of agreement about measuring either concept, it is not surprising that the problem of linking access and health status has posed a considerable challenge.

After an introductory discussion on the concept of health status and its measurement, this paper critically reviews existing measures of access that are linked to health status. As part of this task we explore the underlying assumption that changes in levels of access are likely to result in changes in health status. We then provide recommendations as to how the existing measures might be modified and expanded to provide an improved answer to the question: Given the health status reported by those in a survey, is their level of utilizing medical care services appropriate? Our focus is limited to those measures that can potentially be put to regular use during the next decade in household survey samples of the U.S. population.

From the Departments of *Medicine, †Health Services, and ‡Biostatistics, University of Washington, Seattle, Washington.

This study, which was originally prepared for the Commission, is reprinted with permission from Medical Care, Vol. 19, No. 12 (Supplement), 1981.

February 1981

Figure 1 presents a framework for our discussion of the relationship between health status and access to medical care. It depicts access to the medical care system as just one of several determinants of health status. A major task for health policy is to promote equitable access to those medical services that will optimize the health status of individuals in a population.

Health Status Measurement

To discuss health status measurement, it is first necessary to define the concept of health that one desires to measure. Generally, health is viewed as a multidimensional construct. The World Health Organization defines it as an optimal level of three separate dimensions: physical, mental, and social well-being.[1] The impact of medical care on the different dimensions may vary, and it may be necessary to assign weights to each of them for policy-making purposes concerning access to medical care. Using hypertension as an example, it is generally agreed that treatment is effective in prolonging life and avoiding major morbid events.[2,3] Yet in contradistinction to this important positive effect on physical health, there is some evidence that for many individuals a labeling effect occurs that leads to greater absenteeism from work—a negative effect on social health.[4] To date, most large-scale research on health status vis-a-vis access to care has assigned highest priority to the physical dimension

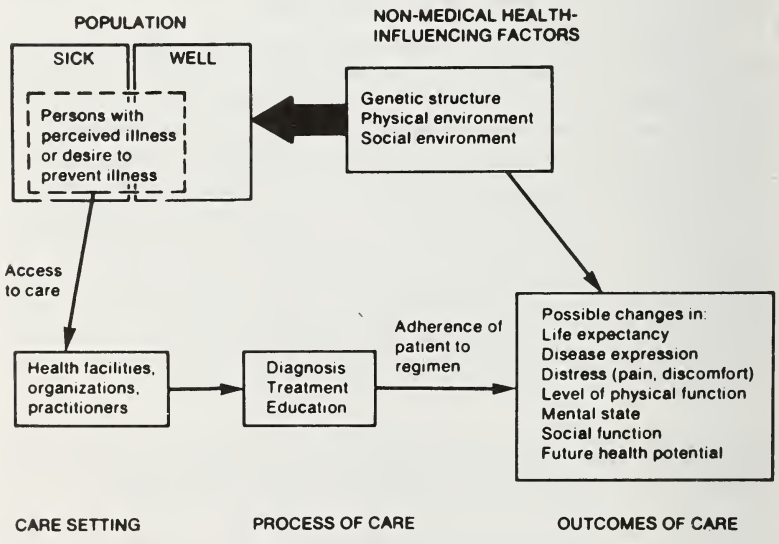


Figure 1. Model of a single episode of medical care. (From INTRODUCTION TO HEALTH SERVICES, edited by Williams SJ, Torrens PR, and printed with permission of John Wiley and Sons.)

of health; modest priority to the mental dimensions; and lowest priority to the social dimension. We now consider the three dimensions separately to determine how well each is measured in routine surveys.

What is physical health? Conceptually, it can be viewed as a combination of physiologic health (physical state and function of biologic organ systems) and functional status. Physiologic health includes presence or absence of commonly defined diseases; the ability of various organs to function (i.e., pulmonary function testing); and biochemical measurements reflecting potential for future disease (i.e., blood cholesterol determinations). Complementary measures of physiologic health might include items such as age, family history of disease, and health habits such as smoking; these factors can contribute to estimates of an individual's future potential physiologic state.

For policy purposes, a question arises concerning the ability of survey questions to approximate results of studies of physiologic health that use more direct operational tools, such as physical examination, physiologic and biochemical tests. When this has been studied, it appears that, in general, surveys underestimate the frequency of derangements of physiologic health. If we focus mainly on those common conditions for which care is efficacious, there is evidence that surveys can yield satisfactory estimates for planning purposes.[5]

Turning to functional status as a measure of physical health, there is much in the literature on the use of survey instruments. A review by Stewart et al. identified six categories used by various groups: self-care activities, mobility (i.e., limitations), activity (i.e., running), role activities, household activities, leisure activities.[6] These comprehensive, but not mutually exclusive, categories can all be measured with survey interviews, and considerable information is available concerning their reliability and validity.[7,8]

In considering mental health measurement, it is important to note that a significant portion of medical care services is directed at improving mental health. Accordingly, it is appropriate to include measures of mental health in policy-oriented studies of access to care. Ware et al. have reviewed the most commonly applied instruments in general population surveys.[9] They identified four constructs that appear to be reliably reported in surveys: anxiety, depression, positive well-being, self-control. However, uncommon but important psychiatric disorders such as manic-depressive illness, and behaviorally manifested disorders such as alcoholism may not be reflected adequately in surveys.

Measures of social health have not received extensive testing in large-scale health surveys. The review of Donald et al. identified four general areas of social function: family and

home, work or major role activity, personal social life outside the home, and community activities.[10] These categories appear more controversial in terms of their relationship to personal medical care services, but their overlap with the categories assessed by physical and mental health suggests the appropriateness of their inclusion as measures of health status.

Beyond the somewhat specific item content in measures of physical, mental, and social health, measures of general health perceptions have also been constructed. Their value stems from their ability to allow the respondent to self-weigh the importance of the dimensions of health. A recently designed measure assesses general health perceptions of past health; current health, health outlook; health worry; resistance/susceptibility; and sickness orientation.[11] In preliminary testing, responses to these general health perception measures are not entirely predicted by the separate measurement of physical, mental, and social health.

Considerable household interview data are available on health status, although for the most part they are based on measures other than the comprehensive constructs that we have briefly reviewed. The National Health Interview Survey from the National Center for Health Statistics (NCHS) and the periodic surveys performed by the Center for Health Administration Studies and the National Opinion Research Center (CHAS-NORC) provide information on physical health and general health perceptions. The measures in these surveys are supplemented by those in additional routine surveys such as the U.S. Immunization Survey (Center for Disease Control). The Health and Nutrition Examination Survey (NCHS) has gathered national data on depressive symptomatology. Unfortunately, there is only a small body of population-based survey data on social health available, although information from the prospective studies of health in Alameda County is a notable exception.[12] The Rand Corporation Health Insurance Study has put to use comprehensive constructs to gather health status data for each dimension of health.[8]

Results from these surveys and others point to continued variation in health status when populations are aggregated by sociodemographic variables of policy interest such as age, place of residence, region, income, and race. It is beyond the scope of this paper to review these variations in health status; nor do we plan to describe differences in health-service utilization, except to state that access has increased in recent years among those groups with a traditionally lower access. Considerable recent work has been done in these areas, including that which is described in the Aday and Andersen paper in this supplement and in a recent review by Davis et al.[13-21]

Nonetheless a question remains: Is access now adequate for groups such as the poor, given the increasing similarity between

their utilization rates and those of other groups? The incorporation of utilization levels and health status into a single measure, as attempted by several researchers, appears appropriate with respect to the question of equity of access. It is less clear whether the measures are adequate as linked measures of appropriate utilization necessary to optimize health status. At this point we review what is known about the effect of changes in utilization of care on the health status of populations.

The Link Between Health Status and Access

Most data available for review are indirect in that they do not come from studies specifically designed to measure the effect on health status of increases in utilization. At the national level, measures of longevity have shown continued increases at a time when utilization of health resources has increased. This is particularly true for groups previously having less access—the elderly and minority groups,[15,16,22] Infant mortality and age-adjusted death rates for many diseases, including the chronic diseases, have declined. These trends are encouraging and we believe that the provision of medical services has had a positive influence.

In terms of physiologic health, evidence exists to support the use of a series of immunizations in the epidemiologic control of certain infectious diseases. Tuberculosis and venereal disease screening is efficacious, especially in high-risk populations.[23] Early detection through screening of cancers of the cervix, rectum, and breast appear to improve longevity.[24] A major indicator of improved physiologic health that has been linked to personal medical care is the proportion of persons with hypertension who are well controlled and for whom evidence suggests a decreased risk of stroke and heart disease, as well as increased longevity.[3,15,16] At the national level, this measure has shown considerable improvement as access to care has increased.

Turning from physiologic health to physical functioning, it appears that as access to care increases so do the number of disability days and the number of symptoms experienced. Thus, at the national level, although there has been a higher use of health services among the poor and nonwhite, there has also been an accompanying increase in restricted activity days due to chronic conditions.[17] In a study designed to address more directly the relationship between access and health status, Diehr et al. have shown that in a low-income but nonwelfare population, increased access to care was accompanied by both increases in disability days and in the number of symptoms reported.[25] The increased reporting of symptoms suggests that the overall decrease in health status could not be attributed solely to disability that was due to physicians' recommendations for limitation of activity. Whether the increased reporting

of symptoms was due to increased awareness of the importance of certain symptoms or to a negative effect of labeling cannot be ascertained. Also, it is not known whether these findings would be duplicated if other, more comprehensive measures of functioning were studied in place of disability days.

Many reasons for a negative association between increased access to care and physical function, measured by disability days and reporting of symptoms, can be posited. Possible explanations include any combination of these points:

(1) Increased access to care has an actual potential for improving health, but the kind of care available to some population groups, such as the elderly, is of poor quality and/or poorly matched to their needs.

(2) Although access has improved relative to need, the levels of access are still insufficient to positively affect disability.

(3) Because of social, environmental, and occupational factors, disability is increasing at a rate that negates the positive effect of increases in access.

(4) Increased access results in increased detection of disease and thus a sick role is appropriately legitimized.

(5) A biologic explanation might be that increased access has been accompanied by increased longevity. The people who otherwise would have died are more disabled and/or have some treatment-related side effects. The survivors contribute to a larger pool of disabled. If this explanation is correct, then usual measures of disability might use the number of people in a birth cohort as the denominator rather than number of people currently alive.

(6) Increased access results in a short-term "catch-up" for care such as elective surgical procedures. This results in more short-run disability, but may lead to improved health over the long run.

(7) Increased access results in increased chances of well people being explicitly or implicitly labeled as sick and unnecessarily adopting a sick role.

(8) Increased medical care actually causes more physical harm than good.

Although we do not attempt here to tease out which explanations are most valid, it is clear that for policy purposes most measures of access that are linked to measures of health status should have a net positive relationship. Otherwise, if measures of health status deteriorate, a cycle is established in which decreased health status triggers more access, which diminishes health status! With these concerns in mind, we can now proceed to a discussion of specific measures that purportedly allow one to examine the appropriateness of access based on the reported health status of groups.

Measures of Access Based on Health Status

Use-Disability Ratio. The use-disability ratio (U-D) was designed to measure the utilization of services in the context of the need for those services. The measure employs days of disability as a proxy measure of need for care among different populations. The U-D ratio is defined as the ratio of physician visits to the days of disability reported in a given period. As used by Aday and Andersen, it is computed by dividing the mean number of physician visits in a given year by the number of disability days in the same year and multiplying the result by 100.[19] For each individual in the survey, physician visits are his or her own estimate of the number of visits experienced in the previous year. Likewise, disability days are determined by asking how many days altogether the person had to stay in bed or cut down on the things he or she usually did during the year because of illness or injury. Limiting the ratio to persons who reported no more than 100 disability days helps insure that a few severely disabled people do not skew the population estimate downward.

The U-D ratio may be a useful measure of social equity, but due to certain methodological and conceptual limitations it is probably not a useful measure of medical needs. From a methodological perspective, it can be criticized because it is imprecise. Both the numerator and denominator depend on patient recall. Although population-based estimates of recall bias for aggregate physician visits have been developed, they are not applicable to individual patients. The U-D ratio is constructed on a patient-by-patient basis and then it is aggregated for the population. Because recall bias has not been conjointly studied for both disability days and physician visits for a given patient, it is not known whether there is a similar bias in the numerator and denominator for different patient groups. Differential recall bias in the numerator and denominator could significantly distort estimates of need measured by the U-D ratio.

A second methodological concern is that the physician visits used in the numerator need not be linked to the episodes of disability in the denominator. Because there is a markedly skewed distribution of disability among individuals within subpopulations, the population-based estimates of disability do not reflect the burden borne by the small subpopulation of greatest interest.[18]

In addition to the methodological concerns, there are several conceptual shortcomings to the U-D ratio as a measure of need. First, the ratio is "best" when a physician visit occurs for each disability day. This problem certainly raises concern about unnecessary care. It could be resolved by using a measure that links an initial visit to an episode of disability (this is discussed in detail below). Another concern is that respondents are asked about illnesses that restrict usual activity, including work days lost. If work loss is used, there is a confounding factor of social

contributions to work loss that is independent of medical care. Because a specified physical disability is more likely to cause work loss for persons doing physical labor, the ratio may understate appropriate use of physician care for blue collar groups compared with the ratio for white collar workers. This problem could also be resolved by changing to a measure based on an episode of disability rather than total disability days.

Symptoms-Response Ratio. In contrast to the U-D ratio, the symptoms-response ratio (S-R ratio) has greater validity as a measure of appropriate access. The S-R ratio directly relates physician utilization to the existence of one or more symptom states for which a panel of physicians determines whether a visit should be made. The estimate of an appropriate response is obtained by having the panel of physicians estimate the proportion of people with a specific symptom complex who *should* see a doctor. Thus, it is an attempt to ascertain an appropriate level of utilization for a given health state. This measure was developed by Andersen and the ratio is calculated as[26]:

$$\text{Symptoms} - \text{Response Ratio} = \frac{A - E}{E}(100)$$

where A = actual number of persons who contacted a doctor at least once for symptom, and

E = physicians' estimate of number of persons who should contact a doctor for symptoms.

A positive value for the ratio implies that more people contacted a doctor than the panel of physicians thought necessary, and a negative number means that fewer contacts were made than were recommended by the panel.

Despite its conceptual attractiveness, the S-R ratio does have significant limitations as an attempt to link health state (occurrence of selected symptoms in the prior year) to an estimate of appropriate use of medical care. One major problem arises from our lack of knowledge concerning the relationships among the presence of a particular symptom, presence of a treatable condition, and the efficacy of treatment. If care is ineffective from a biologic perspective, no net improvement in physical health state is likely to occur with a given level of response to selected symptoms. Even when care might be theoretically effective for selected individuals with a given symptom, the net burden of misdiagnosis or mistreatment of the remaining individuals may mean a reduction in the health status of the population with that symptom. This effect might occur, for instance, in the case of persons with frequent headaches in a particular age group. Physicians could estimate that 30% of such individuals should see a doctor for this symptom. In this group, tension headache will be the most common diag-

nosis made. Many patients therefore, will be diagnosed as having either anxiety or depression, and a number will be treated with various psychotropic agents—some of which will be appropriate, others not.

The net outcome of potential mistreatment for the original population base cannot be determined from existing data, but it is possible that the health status of the overall group (using a broad-based measure) might be worse rather than better. In essence, the likelihood that health status will be improved is a function of the probability that a correct diagnosis is made, a correct treatment given, and the relative efficacy of the treatment. Whether this postulated result would occur using policies designed to improve access based on the current S-R ratio is questionable. It is clear, however, that a much better understanding of the net impact of a care system on all patients presenting with a given symptom is needed to optimally use the S-R ration as a measure of access based on need.

Evidence that physicians may be poor judges of an appropriate response to symptoms casts further concern about the symptoms-response ratio as currently constructed. For instance, Berg and LoGerfo reviewed illness diary data from a longitudinal study of families enrolled in a prepaid group practice.[27] The families' responses to symptom complexes were compared with the recommended action suggested in a book specifically designed to minimize physician visits. The surprising results suggested that had patients followed the book's recommendation, visits would have increased above the actual response. When the illness outcomes were analyzed, no difference in illness could be detected between those who called a physician and those who did not, despite the similar symptom complexes.

A further methodologic concern with the S-R ratio is the relative infrequenty of symptom complexes in a particular individual. For instance, in a sample of low-income, non-Medicaid patients in Seattle, 54% reported none of the symptom complexes.[28] Moreover, most of the reported symptom complexes were in clusters for which only a small per cent are predicted to "need" to see a physician. Thus, as currently constructed, the level of expected response is low, requiring large samples in order to have discriminative power. There is an additional problem with the S-R ratio that is derived from the fact that all symptom clusters are counted rather than one cluster per patient. This means that a patient with multiple symptom clusters will weight the estimates of access differently than a patient with only one cluster. The resulting estimate of access then relates more to illnesses than to individuals with illness. In the Seattle sample, 25% of people had two or more relevant symptom clusters in the previous year.[28]

Having dealt with some imperfections of the S-R ratio for those aspects of health for which it could be useful, we now turn to the problem for which it is presently not applicable—care for those with asymptomatic disease and for those who are healthy. For well people, the S-R ratio fails to measure appropriateness of access to preventive services. Because there is little disagreement about the need for several selected preventive medical services, a minimum estimate of appropriateness of visits can be developed and combined with the S-R ratio in a nonarithmetic fashion. Also, there are asymptomatic diseases that require regular care once detected and of which the patient might be aware as a result of prior medical contacts. Paramount among these is hypertension. Because this is such a common disease among adults, especially in target groups of policy interest, a disease-response category should be added to the S-R ratio if it is to be used as a keystone of policy research.

Episode of Illness Analysis. Having seen that the U-D ratio is not useful for making judgments about appropriate utilization and that the S-R ratio is an incomplete and professionally derived measure of need, we turn now to a measurement strategy that incorporates greater elements of patient demand for services—the episode of illness analysis. As originally conceived, the episode of illness approach was designed to provide richer information about initial health care and follow-up in response to an individual's needs.[29] In order to collect the data, individuals are asked whether they were ill during a specified period, and whether this caused a certain number of disability days. Then, for a period of illness, the respondents are asked a series of questions about whether they sought care, the nature of the care sought, and the perceived seriousness of the problem.

Aday et al. report an expanded approach that includes questions on the current status (outcome) of the problem.[19] They categorize episodes of illness into four types (Table 1). The value of this typology is that it allows one to analyze access on the basis of professionally estimated need for care and patient worry, both of which influence the course of an illness.

Table 1: Illness Types

Type	Need for Care (Physician Judgment)	Respondent Worry Level
I	Elective	Little
II	Elective	Considerable
III	Mandatory	Little
IV	Mandatory	Considerable

Analysis of the information obtained through this approach includes a display of the proportion of individuals contacting a physician, the proportion referred for consultation, an index of continuity of care, and the outcome of the inciting illness. Aday and Andersen applied the four episodes of illness types to their 1976 CHAS-NORC sample population.[20] They analyzed those episodes of illness in the prior year that had caused either three days of disability or considerable pain and worry. Virtually no differences were found in the per cent of individuals who contacted a physician for episodes of illness among various subpopulations categorized by age or place of residence or income. Continuity, need for hospitalization, and the percentage of patients not cured did vary among the subpopulation categories. It is not known whether these differences reflect cultural peculiarities in illness perception and response, quality of care factors, or differences in disability related to social factors.

Although the episode of illness approach is attractive in many ways, there are some problems. Perhaps the greatest single difficulty is that the method is cumbersome for patients who have complex illnesses because it traces all subsequent contacts for that illness. The extra interview time involved might be worthwhile if significant differences among population groups are found in the analysis. The two previously cited studies, however, point to frequently small differences in follow-up once adjustment is made for illness levels and for whether an initial contact is made.[19,29] Furthermore, although continuity is an important dimension of care, especially as it relates to satisfaction, there is no convincing data that marginal increases in continuity result in better overall outcomes. Thus, when examining episodes of illness, further work is needed to decide whether, from an efficiency perspective, the increase in interviewers and staff time necessary for continuity analysis is worth the effort.

One additional issue in regard to the episode of illness method is worthy of discussion. No summary index of appropriate response has been constructed. To do this, the data on outcome status and on health care sought should be combined so that a summary index can be developed. Figure 2 displays the individual classifications of each illness episode by illness type, the outcome, and whether a physician was seen. As an example of how a summary index might work, a patient with episodes for which care is elective and causes little worry (Type I) is considered as having adequate access (boxes 1+, 1−, 2+, 2−, 3+, 3−, 4+, 4−). Access for Type II illness is adequate if a physician was consulted (boxes 5+, 6+, 7+, 8+) or if no physician were consulted but the illness was cured or improved (boxes 5−, 6−). If the illness did not improve and there was no physician visit, access could be viewed as inadequate even though care was elective (boxes 7−, 8−). Cured Type III illness

would be considered having adequate care with or without an associated visit (9+, 9−), as would less-then-cured Type III illness with an associated visit (boxes 10+, 11+, 12+). Type III illness that was not cured and for which there was no visit is considered inadequate access (boxes 10−, 11−, 12−). The adequacy of access for Type IV illness is as for Type III illness, except that because of the degree of patient worry, even the cured illness should be associated with a visit (box 13−).

A summary index such as the one outlined above may need considerable development, but it is particularly attractive because inclusion of outcomes in the index allows links to be made to the quality of care. Neither the S-R ratio nor the U-D ratio can do this.

Recommendations for a Comprehensive Measure

We now outline the scope of services that should be included in a more comprehensive construct designed to assess access as related to health status. We emphasize that these recommendations focus primarily on the question: What measures reflect a level of access that is likely to optimize health status? Furthermore, we have attempted to take into account the existing scope and utilization of medical services that exist in the United States as well as that which is likely to occur over the next decade.

Our first recommendation is that an index of access to a multitude of different types of services be designed for analysis in a disaggregated profile rather than as an overall summary. Just as there is more than one dimension to health status, medical services are also heterogeneous. Specific services can be viewed as appropriate for preventive, acute, or chronic modalities. Indeed, much of the literature on access has already put to use measures that reflect access to these different types of services. Valuable information would be lost if both health status and services were aggregated into a single summary value.

Preventive Care. To assure optimal benefits with respect to future health status, an access measure should be linked to an overall preventive service index that is age and sex specific and recognizes special at-risk groups. Thus, if such a measure includes questions concerning detection of high blood pressure, it should expect higher use for black and older populations (over 40 years). Conversely, measures of access to Pap smears should be adjusted for the frequency of women who have undergone a hysterectomy (over 20% of women above age 40). Numerous similar restrictions and qualifications are in order. Overall, the preventive services measures included should be those for which consensus exists concerning efficacy.

Acute Care. The existing use-disability ratio should not be used as a measure of access to care because of its nonspecificity.

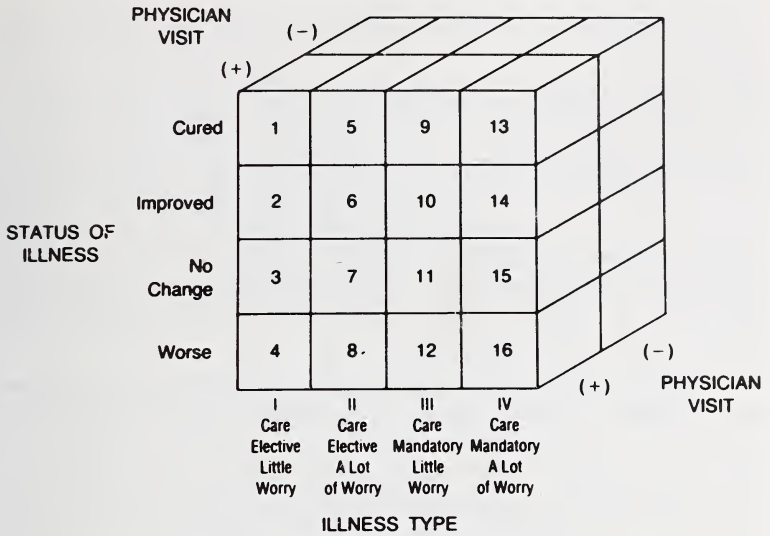


Figure 2. Framework for the development of a summary index based on episodes of illness.

For acute symptomatic illness we recommend that an access index be episode-linked and include degree of patient worry as a component. To adjust for self-limited disease, a measure of current symptom status should be included in a manner analogous to that described earlier in the episode of illness critique. An index that includes only initial contact and is linked directly to those people in a population who experience the illness, represents substantial improvement over the U-D ratio. This notion of measuring only the initial contact assumes that subsequent contacts recommended by physicians are likely to occur for patients with serious illness.

Chronic Disease Care. For chronic disease and disability, an access index should include diseases of which the patient is aware and for which care is efficacious. For instance, a specified minimum number of visits should be expected for all people who have hypertension, diabetes, or currently symptomatic arthritis. Adjustments should be made for the presence of more than one disease in a given person.

Hospital Care. Hospital services can be viewed as either acute and urgent or elective. At this time elective care is more likely to be affected by policy concerning access to care. Access to elective surgical care should include an index that ascertains whether people have been told they should have an operation for a condition that is currently symptomatic, and for which they have not received care because of a specific financial or geographic barrier.

An overriding concern for all types of care must deal with inherent shortcomings of household surveys themselves. Recent studies of the National Health Interview Survey sample of the U.S. population, for example, have shown that some groups of greatest policy interest (i.e., minorities, migrant farm workers) are not interviewed in adequate numbers.[30,31] Furthermore, Sparer and Okada have described the lack of correspondence between national data and findings at the local level for chronic conditions rates.[32] Solutions to these problems have been proposed, the most important of which is likely to be the supplementation of national survey information with local data on specific population subgroups.

In addition to the above recommendations, we emphasize that further research is needed to address the relationships among health status, access to care, and subsequent outcomes. Given the generally high proportion of people who receive at least some care in a given year, greater emphasis should be on the nature of the services received and outcomes of care for various populations. As a beginning step, it would be helpful to have national and geographic estimates of access and health status routinely standardized to a given population base for greater precision in estimating trends over time.

Substantial research is needed to explore what mix of medical, social, and vocational services will lead to the greatest reduction in disability that is due to various diseases once they have occurred.

Finally, given current economic and social trends, it is quite likely that access to care for low-income people will be decreased and/or further segregated from care for higher-income groups. If so, high priority should be given to establishing a comprehensive, national research strategy to monitor changes in health status by using measures of longevity, physiologic health, physical function, and mental health. These changes should be carefully assessed in the context of changes in access to care.

Acknowledgments

The authors thank Drs. Robert Brook, John Ware, and colleagues at the Rand Corporation for comments on this manuscript, and for providing the unpublished information from the Health Insurance Study.

References

[1] World Health Organization. BASIC DOCUMENTS, 26th ed. Geneva: World Health Organization, 1976.

[2] Veterans Administration Cooperative Study Group on Antihypertensive Agents. *Effects of treatment on morbidity in hypertension.* JAMA 1976;202:116.

[3] Hypertension Detection and Follow-up Program Cooperative Group. *Five-year findings of the hypertension detection and follow-up program.* JAMA 1979;242:2562.

[4] Haynes RB, et al. *Increased absenteeism from work after detection and labeling of hypertension patients.* N Engl J Med 1978;299:714.

[5] Brook RH, et al. Conceptualization and Measurement of Physiologic Health for Adults. Hypertension, R-2262/3-HHS. Santa Monica, CA: The Rand Corporation, 1980.

[6] Steward AL, et al. Conceptualization and Measurement of Health for Adults in the Health Insurance Study, Vol. II. Physical Health in Terms of Functioning, R-1987/2-HEW, Santa Monica, CA: The Rand Corporation, 1978.

[7] Ware JE, et al. Conceptualization and Measurement of Health for Adults in the Health Insurance Study, Vol. I. Model of Health and Methodology, R-1987/1-HEW. Santa Monica, CA: The Rand Corporation, 1980.

[8] Brook RH, et al. *Overview of adult health status measures fielded in Rand's health insurance study.* Med Care 1979;17:Supplement.

[9] Ware, JE, et al. Conceptualization and Measurement of Health for Adults in the Health Insurance Study, Vol. III. Mental Health, R-1987/3-HEW. Santa Monica, CA: The Rand Corporation, 1979.

[10] Donald CA, et al. Conceptualization and Measurement of Health for Adults in the Health Insurance Study, Vol. IV. Social Health, R-1987/4-HEW. Santa Monica, CA: The Rand Corporation, 1978.

[11] Ware JE, et al. Conceptualization and Measurement of Health for Adults in Health Insurance Study, Vol. V. General Health Perceptions, R-1987/5-HEW. Santa Monica, CA: The Rand Corporation, 1978.

[12] Berkman LF, Symes SL. *Social networks, host resistance and mortality: a nine-year follow-up study of Alameda County residents.* Am J Epidemiol 1979;109:186.

[13] United States Department of Health and Human Services: *Health of the Disadvantaged. Chartbook II, 1980.* DHHS Publication No. (HRA)80-633.

[14] Lerner M, Stutz R. *Have we narrowed the gaps between the poor and the nonpoor?* Med Care 1977;15:620.

[15] Soltero, I, et al. *Trends in mortality for cerebrovascular diseases in the United States, 1960-1975.* Stroke 1978:9:549.

[16] Cooper R, et al. *The decline in mortality from coronary health disease, USA 1968-1975.* J Chronic Dis 1978;31:709.

[17] Wilson R, White E. *Changes in morbidity, disability, and utilization differentials between the poor and nonpoor: data from the Health Interview Survey: 1964 and 1973.* Med Care 1977:15:636.

[18] Newacheck PE, et al. *Income and illness.* Med Care 1980;18:1165.

[19] Aday LA, Andersen R, Fleming GV. Health Care in the US: Equitable for Whom? Beverly Hills, CA: Sage Publications, Inc, 1980.

[20] Aday LA, Andersen R. *Equity of access to medical care, a conceptual and empirical overview.* Med Care (Supplement) 1981;19:4.

[21] Davis K, et al. *Access to health care for the poor: Does the gap remain?* Ann Rev Public Health 1981;2:159.

[22] Rosenwaike I, et al. *The recent decline in mortalityof the extreme aged: An analysis of statistical data.* Am J Public Health 1980;70:1074.

[23] Canadian Task Force. *Periodic health examination.* Can Med Assoc J 1979;121:1193.

[24] Eddy D. *Guidelines for the cancer related checkup: recommendations and rationale.* CA-A Cancer Journal for Clinicians 1980;30:194.

[25] Diehr PK, et al. *Increased access to medical care, the impact on health.* Med Care 1979;17:1.

[26] Andersen R. *Health status indices and access to medical care.* Am J Public Health 1978;68:458.

[27] Berg A, LoGerfo J. *Potential effects of self-care algorithms on the number of physician visits.* N Eng J Med 1979;300:535.

[28] Diehr PK, et al. The Seattle Prepaid Health Care Project: Comparison of Health Services Delivery, Chapter II. N.T.I.S. Publication #(P.B.267-490), 1976.

[29] Richardson WC. *Measuring the urban poor's use of physician services in response to illness episodes.* Med Care 1970;8:132.

[30] Salber EJ, Beza AG. *The health interview survey and minority health.* Med Care 1980;18:319.

[31] Rice DP, et al. *Household health interviews and minority health, the NCHS perspective.* Med Care 1980;18:327.

[32] Sparer BA, Okada BS. *Chronic condition and physician use patterns in ten urban poverty areas.* Med Care 1974;12:549.

Concepts of Medical Underservice: A Review and Critique

Eugenia S. Carpenter, M.P.P.*

Introduction

The concept of medical underservice implies that a geographic area or a specific population group is denied access to medical care because of geographic, physical, financial, social, or cultural barriers. This paper examines the various approaches taken by agencies at the Federal and state level, as well as by such private entities as foundations and professional associations, to define and measure medical underservice. Why are those definitions and measurements important? Primarily because governmental and private entities use them to allocate finite health care resources to improve access.

Medical underservice is a normatively derived concept and depends on two underlying assumptions: (1) Some basic right to health care inures to all citizens; (2) there is a direct relationship between the availability of health care resources (personnel, facilities) and their utilization and a positive relationship between utilization and health care outcomes. Like other consensual goals, the right to health care is susceptible to a wide range of operational definitions. What level or amount of health care is "reasonable" or "adequate"? What range of health care services should be available and accessible—preventive, curative, "caring," psychiatric?

Despite a lack of agreement on answers to these questions, Federal and state governments over the past two decades have developed a wide range of policies and programs aimed at

*Department of Health Planning and Administration, School of Public Health, University of Michigan, Ann Arbor, Michigan. Research on which this paper is based was supported by a grant from the W.K. Kellogg Foundation.

January 1982

removing "racial, economic, social, or geographic barriers which now prevent many of our citizens from maintaining adequate health care protection."[1] These policies have targeted remedial efforts to those geographic areas and population groups with inadequate access to health care. In targeting program resources, inequality of access has been described with such terms as health manpower shortage areas, medically underserved areas, and structurally underserved populations. The lack of precision of these different terms is evident when one examines the operational criteria and the measurements employed in defining medical underservice.

Current definitions of medical underservice reflect differences in (1) perceptions of the nature of the problem of unequal access; (2) objectives of policies and programs to remove such inequities; and (3) philosophical approaches to a just allocation of health care resources. Perceiving underservice as due to a lack of available health care resources—physicians, hospitals, clinics—leads to a focus on geographic maldistribution of resources and on such factors as travel time, population densities, and other spatial considerations. If underservice is seen as deriving principally from other obstacles then broader issues of access will arise and the focus will include economic, social, and demographic factors. However, even if agreement were possible on the meaning and operational measurement of such terms as access and basic needs, approaches that seek to minimize disparities among geographic areas and population groups would differ philosophically from those seeking to assure a basic "floor" of health care protection to all. In reviewing the operational definitions of "underservice" or "shortage" in current usage, we find that both approaches may be included in the same index.

Components of Shortage or Underservice Measures

Efforts to define and quantify concepts of medical underservice or health provider shortages have generally incorporated measures of such components as availability, access, need, utilization, and rational service areas. These components have been defined in terms of supply of health care providers, utilization of services, health status measurements, and socioeconomic indicators. The difficulty of establishing uniform definitions for these concepts is compounded by the fact that the same indicator, e.g., utilization, may be used to measure more than one component.

Availability. Availability is most often expressed as the supply of health care resources, personnel, and facilities in a

[1] Nixon, Richard M. PRESIDENT'S HEALTH MESSAGE OF 1971, Washington, D.C.: The White House, Feb. 18, 1971.

given geographical area or population. Bed:population ratios and physician:population ratios are among the oldest tools employed by health planners to measure adequacy.

Although a widely used criterion for medical underservice, the physician:population ratio is ambiguous as to what it measures and what it means. Four methods of estimating physician manpower requirements have been developed: medical need (epidemiological), professional judgment, demand/productivity, and HMO prepaid practice. Over a 50-year period in the United States, estimates of overall physician requirements, based on a variety of assumptions and methodologies, have ranged from 100-192 physicians per 100,000 population.[2] Many of these estimates represented arbitrary norms derived from empirical measures of "favorable" supply or "appropriate" utilization. Epidemiological models, exemplified by the work of the 1930 Committee on the Costs of Medical Care and more recently the Graduate Medical Education National Advisory Council (GMENAC)[3] are perhaps closest to matching physician supply to medical need. Even this approach leaves a large area for subjective judgment about treatment modalities, as well as about the substitutability of different types of health care personnel (among physician specialties and between physician and nonphysician providers) in producing units of treatment.[4] In short, the variety of assumptions about practice norms and health care organization underlying different manpower models results in a considerable range of estimates of the "appropriate" physician:population ratio.

---

[2] See, for example, Mountin, J.W., *et al.* Health Service Areas: Estimates of Future Physician Requirements. Public Health Bulletin No. 305, Washington, D.C., U.S. Government Printing Office, 1949. *Physicians for a Growing America*, report of the Surgeon General's Consultant Group on Medical Education. PHS Publication No. 709, Washington, D.C., U.S. Government Printing Office, 1959. Report of the National Advisory Commission on Health Manpower, Vol. 2, Appendix 6. Washington, D.C.: U.S. Government Printing Office, November 1967. Tokuhata, George K., *et al. Health Manpower Distrubution in Pennsylvania.* Am. J. Public Health, 65, Aug. 1975, 837-848. Kansas Medically Underserved Areas 1980. Prepared by Office of Health Planning and Bureau of Registration and Health Statistics, Kansas Department of Health and Environment, Topeka, Kansas, Dec. 31, 1980.

[3] The Committee on the Costs of Medical Care. The Fundamentals of Good Medical Care. Chicago: University of Chicago Press, 1933. Summary Report of the Graduate Medical Education National Advisory Committee, Vol. 1. Office of Graduate Medical Education, HRA, PHS, USDHHS, Sept. 1980.

[4] See, for example: Mason, Henry P. *Manpower Needs by Specialty.* J.A.M.A. 219:1621-1626, March 20, 1972. Schonfeld, Hyman K., Jean F. Heston, and Isidore S. Falk. *Numbers of Physicians Required for Primary Medical Care.* New Eng. J. Med., 286:571-76, March 16, 1972. Stevens, Carl M. *Physician Supply and National Care Goals.* Industrial Relations, 10:119-44, May 1969.

Access. "Access" is widely used, yet remains an ill-defined term in health services literature. Aday and Andersen stated that the concept of access is "more of a political than an operational idea."[5] Parker's working definition of "accessibility" is "the ability to reach, obtain, or afford entrance to service."[6] As applied to measures of medical underservice, access has most often been measured in terms of distance and travel time. The Health Manpower Shortage Area (HMSA) criteria for underserved populations recognize refusal to accept certain types of patients or Medicaid reimbursement as a barrier to access. The California physician shortage area designation takes account of cultural barriers in making special provisions for ethnic minorities and non-English-speaking populations. Some treatments of the concept of access have attempted to measure it by outcomes, specifically in terms of utilization of services. In this review, the term will be used more narrowly in the sense of accessibility, with utilization measures discussed under need and demand.

The spatial dimension of access can be represented by distance and travel time to sources of care. National and subnational studies indicate that most of the population is within reasonable travel time of regular sources of medical care and of hospital service.[7,8] Exceptions are found among the poor, the elderly and nonwhites.

In general, 30-minute travel time has become accepted in the health planning community and has been incorporated as a standard in several health planning guidelines. Except for emergency medical services, there is little empirical evidence of the health effects of moderate differences in travel time. Thus, it has been suggested that an emphasis on travel time by itself in the HMSA criteria is not sensible.[9]

The length of time it takes to locate an appropriate source of care, the waiting time for an appointment, and the time spent in the facility, clinic, or physician's office waiting to be seen are

---

[5] Aday, Lu Ann and Ronald Andersen. ACCESS TO MEDICAL CARE. Ann Arbor: Health Administration Press, 1975.

[6] Parker, Alberta W. *The Dimension of Primary Care: Blueprints for Change*. In PRIMARY CARE: WHERE MEDICINE FAILS. Spyros Andreopoulos, Ed. New York: John Wiley and Sons, 1974, 15-17.

[7] Verbrugge, Lois M. *Medical Care of Acute Conditions, United States, 1973-1974*. VITAL AND HEALTH STATISTICS. Series 10, No. 129. DHEW Publication No. (PHS) 79-1557. Hyattsville, Md.: National Center for Health Statistics, Feb. 1979.

[8] Bosanac, Edward M., Rosalind C. Parkinson, and David S. Hall. *Geographic Access to Hospital Care: A 30-Minute Travel Time Standard*. MED. CARE, 14:616-24, July 1976.

[9] Mathematica Policy Research, for Bureau of Health Professions. EVALUATION OF HEALTH MANPOWER SHORTAGE AREA CRITERIA. DHEW Pub No. (HRA) 80-20. U.S. Department of Health, Education and Welfare, PHS, HRA, April 1980.

other measures of access. In an effort to develop a more refined measure of acceptable waiting times, Simon and his colleagues developed an index based on four levels of urgency or severity of illness. Acceptable waits for appointments ranged from one to three hours for critical conditions, one to two days for serious, one to four weeks for chronic, and one or more months for routine care.[10]

Race, ethnicity, knowledge of the health care system, level of education, and ability to communicate are factors that, under given circumstances, can increase the effect of travel distance. The work of Morrill and Earickson and the results of other studies indicate that race, religion, and income operate to influence patient travel behavior.[11] Racial segregation, dual tracks in the health care system (Medicaid vs. private patients), and the perceived differences in quality of physician offices and hospitals tend to modify the effects of distance.

The difficulty of changing and adapting to a new source and place of care is another factor to be considered. Patients may continue to utilize a familiar source of care, even though a change of residence has substantially increased the distance to be traveled. This tendency reinforces the findings of Aday and Andersen about the importance of a regular source of primary care as a component of access.

The literature also suggests that simple measures of time and distance barriers need to be supplemented with more subjective data. Penchansky and Thomas measured access with a survey of patient satisfaction. Defining access as "a measure of the fit between characteristics of providers and health services and characteristics of patients," they focused on patients' perceptions of their ability to locate good health care and their knowledge of how to go about getting needed care.[12] Incorporation of these perceptions of relative ease or difficulty of obtaining care may be needed to provide more sensitive measures of consumers' assessment of access.

---

[10] Simon, Howard, *et al. An Index of Accessibility for Ambulatory Health Services.* MED. CARE 17(9):894-901, Sept. 1979.

[11] Studnicki, James. *The Minimization of Travel Effort as a Delineating Influence for Urban Hospital Service Areas.* INT. J. HEALTH SERVICES 5:4, 1975. Bashshur, R.L., G.W. Shannon, and C.A. Metzner. *Some Ecological Differentials in the Use of Medical Services.* HEALTH SERVICES RESEARCH, Spring 1971:61-75. Morrill, R., and R. Earickson. *Factors Influencing Distances Traveled to Hospitals.* ECONOMIC GEOGRAPHY 46, 1970. Shannon, G.W., R.L. Bashshur, and C.A. Metzner. *The Concept of Distance as a Factor in Accessibility and Utilization of Health Care.* MED. CARE REV. Feb. 1969:143-161. Shannon, G. and A. Dever. HEALTH CARE DELIVERY: SPATIAL PERSPECTIVES. New York: McGraw-Hill, 1974.

[12] Penchansky, Roy and J. William Thomas. *The Concept of Access: Definition and Relationship to Consumer Satisfaction.* MED. CARE 19(2):127-140, Feb. 1981.

Rational Service Areas. The construction of valid measures of supply of health providers depends on the delineation of a geographic unit of analysis. Essentially, the problem is one of identifying market areas in medical services. The term "health service area" or "medical service area," as used in health planning models, denotes a geographic area that is relatively self-contained for the purpose of its residents obtaining specified medical services. That is, for a service area to be "rational" in some normative sense, it must be assumed that few residents move between areas for the purpose of obtaining medical care.[13] Locational analysis, developed principally by economic geographers, provides a framework for examining spatial demand for goods and services. According to central place theory, a hierarchical ordering of economic activities will occur among communities such that there are thresholds of population size and purchasing power for the support of a given mix of economic services.[14] In the most rural areas, a minimal level of economic activity would probably be represented by a general store and a gas station. As with other goods and services, thresholds exist for different levels and types of medical care, ranging from a general or family physician to such highly tertiary services as burn centers and open-heart surgery. The distance that people are willing to travel will vary with the type of good or service. Conversely, population density, income, and travel costs (time and distance) will determine whether a particular good will be produced in an area. This means, for example, that sparsely populated rural areas and areas with mountainous and difficult terrain may not be able to support an economically viable physician practice.

In urban areas, obstacles to defining rational service areas arise from the difficulty of identifying independent and self-sufficient geographic units for small geographic areas. Urban health resources tend to serve populations in overlapping patterns, and the density and diversity of available resources tend to elevate behavioral, cultural, and social factors over geographic ones in determining patterns of access to care.[15] This is not to say that spatial factors, manifested in lack of public transportation and time and in monetary costs of travel, do not constitute formidable barriers to care for the urban poor. But delineation of rational service areas in urban areas requires consideration of economic and social as well as geographic

---

[13] Studnicki, James. *The Geographic Fallacy: Hospital Planning and Spatial Behavior.* Hospital Administration 20:10-21, Summer 1975.
[14] Leyes, John M. *The Delineation of a Hierarchical System of Health Service Areas.* In Proceedings of the Workshop on Health Manpower Shortage Areas, DHEW Pub. No. (HRA) 77-51. U.S. Department of Health, Education and Welfare, PHS/HRA/BHM, Washington, D.C., April 1977, 77-89.
[15] Studnicki, *Geographic Fallacy,* 1975, *op. cit.*

factors, as the work of Berry and his associates has demonstrated.[16] A major criticism of the Federal shortage area criteria is that geopolitical boundaries, usually county lines, are used to delineate medical service areas.[17] On the other hand, use of these boundaries ignores the fact that residents of one county may routinely travel across county lines to obtain goods and services, including medical care. Conversely, in very large and in densely populated counties, viable market areas are likely to be smaller than a county, and the aggregate patterns of health service utilization at a county level will obscure pockets of medical underservice within its boundaries. It must be acknowledged that the use of county and census tract boundaries for identifying medical service areas reflects the availability of demographic and vital statistics data by geopolitical units. In most instances, health service areas so designated are neither "rational" nor empirical.

Need. Implicit in the concept of medical underservice is the notion that some population groups have more than average unmet needs for health services. One approach to determining whether unmet need exists was referred to earlier in discussing the need-based estimates for physician:population ratios. Age-specific incidence and prevalence rates for acute and chronic diseases are used as indicators of the need for care; these measures may be adapted to the characteristics of any given population group to estimate expected incidence rates and the need for medical care services to treat those conditions, following standard professional criteria.[18] In applying the need concept to measurement of medical underservice, the need-based estimate could be compared to actual utilization patterns. Any shortfall might be attributed to insufficient health care services.

A criticism of need-based estimates, whether used to derive physician:population ratios or to develop some normative standards for expected utilization of resources, has been the complexity, difficulty, and costliness of obtaining the data and of making the relevant estimates of resource requirements. For this reason, a number of more simplified, proxy measures of need have been employed, and have been incorporated into some of the measures of medical underservice to indicate a possible lack of adequate levels of care. A frequently cited definition of need for medical services is that of Jeffers *et al.*:

---

[16] Berry, Charles, *et al. Final Report: A Study of the Responses of Canadian Physicians to the Introduction of Universal Medical Care Insurance: The First Five Years in Quebec.* MPR Project Report 78-09. Princeton, N.J.: Mathematica Policy Research, Feb. 1978.

[17] Comptroller General of the United States. HEALTH SERVICE PROGRAM NEEDS ASSESSMENTS FOUND INADEQUATE. HRD 81-63. Washington, D.C.: U.S. General Accounting Office, June 15, 1981.

[18] Roddy, Pamela C. *Need-Based Requirements of Primary Care Physicians.* J.A.M.A. 243(4):355-358, Jan. 25, 1980.

"[The] quantity of medical services which expert medical opinion believes ought to be consumed over a relevant time period in order for its members to remain or become as healthy as is permitted by existing medical knowledge."[19] This definition regards optimum health as the goal of medical services, a somewhat idealistic view in 1982.

Need can be contrasted to a population's wants, commonly derived by surveys of consumers, or to demand for medical services, discussed below. Need indicators of underservice have employed health status indices such as mortality rates and interview reports of symptomatic morbidity as well as clinical reports of morbidity, or indirect proxy measures of need such as a population's demographic characteristics.[20]

Mortality data have the advantage of ready availability, usually in computerized form, and for small geographic areas. For these practical reasons, they are frequently proposed or used as health status indicators. For example, the infant mortality rate is one of the four components of the Federal Index of Medical Underservice. The extensive work of Harris and his colleagues at the Rand Corporation has culminated in planning methodologies based principally on indices of mortality to set priorities for health planning interventions.[21] Other mortality-based indicators include productive years of life lost (PYLL) Index (recommended by Kleinman), excess premature mortality indices, and the index of preventable deaths as a health sentinal event, as proposed by Rutstein.[22]

Indirect measurements of need based on population demographics have included adjustments for the population over 65, female population between 15 and 44 as a proxy for fertility rates, percent of population in poverty status, and other sex, age, and race adjustments to population data incorporated as proxy measures of unmet need for health care. In addition to the indicators cited above, Aday mentions such other measures as low infant birth weight and the incidence of nonmedically attended conditions.[23]

---

[19] Jeffers, James R., *et al. On the Demand Versus Need for Medical Services and the Concept of Shortage.* AM. J. PUBLIC HEALTH, 61:46-63, Jan. 1971.

[20] Aday, Lu Ann. *Criteria for Determining Shortage Areas.* PROCEEDINGS OF THE WORKSHOP ON HEALTH MANPOWER SHORTAGE AREAS, Orlando, Florida, Nov. 8-10, 1976. DHEW Publication No. (HRA) 77-51. Washington, D.C.: U.S. Government Printing Office, 1977, 99-118.

[21] Harris, L. Jeff, *et al.* ALGORITHMS FOR HEALTH PLANNERS: VOL. 1 IN OVERVIEW. Santa Monica, California: The Rand Corporation, Aug. 1977.

[22] Kleinman, Joel C. *Age-Adjusted Mortality Indexes for Small Areas: Applications to Health Planning.* AM. J. PUBLIC HEALTH, 67:834-840, Sept. 1977. Rutstein, D.D., *et al. Measuring the Quality of Medical Care.* NEW ENG. J. MED. 294:588, 1976.

[23] Aday, 1977, *op. cit.*

Utilization patterns, for example, comparing physician visits of population groups in a given period or the proportion of the population with or without a regular source of care, are frequently cited as indicators of unmet need.[24] In their comprehensive analysis of national patterns of utilization, Aday and Andersen found that about 11% of the population had no regular source of care and that these individuals were most likely to fall into specific demographic groups. Not surprisingly, these populations are relatively disadvantaged on other measures of social and economic well-being. Salkever attempted to relate need to the probability of entry into the medical care system, measuring access by the probability of receiving care for a given acute illness episode.[25] Other indices that attempt to combine utilization and need measures include the use:disability ratio, the symptoms:response ratio and the medical severity index.[26]

The use:disability ratio relates consumer perceptions of need as measured by reported in-bed disability days with the actual utilization of physician services. The symptoms:response ratio considers the actual physician visits for a given set of symptoms and the estimates from a panel of medical experts as to what should be the appropriate number of visits for such symptoms. The medical severity index attempts to deal with the problem that a crude comparison of utilization rates inadequately measures; the poor and other groups may have more serious health problems, and therefore they should be expected to have higher utilization rates.

Demand. Health manpower shortages and medical underservice have also been expressed as unmet demand for care. Effective demand for health care services can be directly measured in units of service actually utilized; e.g., physician visits, hospital days. Demand for service units can be converted into derived demand for health manpower. A shortage or underservice would exist where there was a willingness to purchase more units of service than were available. Manpower shortages might be manifested in terms of price increases, long waiting times, substitution of personnel (e.g., physician assistants for physicians), utilization of alternative care modalities (e.g., folk

[24] See, for example: Salber, E.J., *et al. Access to Health Care in a Southern Rural Community.* MED. CARE, 14:971-986, 1976. Greene, S.B., *et al. Distribution of Illness and Its Implications in a Rural Community.* MED. CARE, 16:863-877, 1978.

[26] See Aday and Andersen, 1975; and Davis, K. and R. Reynolds. *The Impact of Medicare and Medicaid on Access to Medical Care.* THE ROLE OF HEALTH INSURANCE IN THE HEALTH SERVICES SECTOR. Edited by Richard N. Rosett. New York: National Bureau of Economic Research, 1976, 391-425. Health Service Research and Training Program. HEALTH SERVICES DATA SYSTEM: THE FAMILY HEALTH SURVEY. Lafayette, Indiana: Health Services Research and Training Program, Purdue University, 1972.

practitioners) or reliance on residents, *locum tenens* physicians, unlicensed physicians, and foreign medical graduates.

Where substantial financial barriers to care exist, as in areas with high concentrations of the poor and the uninsured, effective demand is likely to underestimate medical need as normatively determined. Conversely, where criteria for defining underserviced areas fail to take into account demand factors, the areas so identified may be inappropriate targets for such incentives as start-up subsidies for physicians or loan redemption programs that seek both to attract physicians to underserved areas and to retain them. Policy guidelines for the National Health Service Corps (NHSC), for example, have vacillated between recognizing the appropriateness of permanent subsidies to areas with insufficient demand to sustain viable private practices and emphasizing the Corp's role as "a catalyst for the development of private practice."[27]

The difficulty of relating supply to demand is illustrated by the work of Wennberg and Gittelsohn in studying small area variation in utilization of health care services in New England.[28] Finding that variations in patterns of use of most medical and surgical procedures could not be explained in terms of consumers' age-standardized illness rates or behavior in seeking care, they concluded that patterns of use must be attributable to differences in physicians' clinical decisions. Wennberg and Gittelsohn's findings challenge the validity of econometric, demand-based, and needs-based estimates of health resources requirements, which assume that "current utilization represents an . . . approximation of the optimum situation."

Construction of Comprehensive Measures of Shortage or Underservice

Two principal measures of medical underservice are in use for Federal grant programs: The Health Manpower Shortage Area (HMSA) and the Index of Medical Underservice (IMU), which uses the designation "medically underserved area" (MUA).

Twenty-three programs in five administrative agencies of the Department of Health and Human Services and one program under the jurisdiction of the Department of Agriculture use the HMSA/MUA designation as a basis for resource allocation. Although often the reliance upon the HMSA and MUA designations is intertwined in program administration, the core of programs relying on the HMSA designation differ from those

[27] U.S. House of Representatives, Committee on Appropriations, 1979.
[28] Wennberg, John E. and Alan M. Gittelsohn. A Small Area Approach to the Analysis of Health System Performance. DHHS Pub. No. (HRA) 81-14006 PHS, HRA, BHP. U.S. Department of Health and Human Services, Aug. 1980.

using the MUA designation. Those programs that provide financial aid to health professional students in return for a commitment to practice in an underserved area or medical specialty use the HMSA designation. MUA designations are used by those programs providing some form of support (facility construction, start-up and operating funds, technical assistance) to projects that provide direct health care services. This distinction provides a framework for examination of the different criteria used to define the two designations of medical underservice.

The Index of Medical Underservice (IMU). This index was developed in response to the Health Maintenance Organization Act of 1973 (PL 93-222), which gave priority in funding to the development of HMOs for medically underserved areas. The legislation's definition of medical underservice is "a condition of shortage of personal health services," but the actual development of criteria for the measurement of underservice was left to the discretion of the Secretary of (then) Health, Education, and Welfare (HEW). As developed, the IMU never came to grips with defining medical underservice. Rather, the Health Services Research Group at the University of Wisconsin developed an index made up of a weighted sum of four components: the primary care physician:population ratio, the infant mortality rate, the percent of population under the poverty level, and the percent of the population over 65. As noted above, these components reflect direct or proxy measures of availability and need. Both the authors of the index and others have reported difficulties in applying it to urban underserved areas.[29]

A number of technical and methodological criticisms have been made of details in the IMU.[30] In particular, the question of whether the four components belong together in an index is important, since their combination could produce anomalous results if the values on each of the four indicators were particularly skewed. Davis and Marshall have suggested alternative criteria for developing a medical underservice index including manpower availability, inadequate service utilization, excessive time delays to obtain care, the quality of care, and health status.[31] They recommended an index that would be based upon

[29] Health Services Research Group, Center for Health Systems Research and Analysis, University of Wisconsin. *Development of the Index of Medical Underservice.* HEALTH SERVICES RESEARCH, 10:168-80, Summer 1975.

[30] Wysong, Jere A. *The Index of Medical Underservice: Problems in Meaning, Measurement, and Use.* HEALTH SERVICES RESEARCH, 10:127-35, Summer 1975.

[31] Davis, Karen and Ray Marshall. *Primary Health Care Services for Medically Underserved Populations.* PAPERS ON THE NATIONAL HEALTH GUIDELINES: THE PRIORITIES OF SECTION 1502. DHEW Publication No. (HRA) 77-641. Washington, D.C.: U.S. Government Printing Office, Jan. 1977, 1-23.

objective standards of appropriate care and would measure deviations from these as indicators of medical underservice.

In an empirical test of the IMU, Kleinman and Wilson found that rural MUAs were not markedly different from other rural areas.[32] Kleinman and Wilson essentially agreed with Davis and Marshall in their critique of the IMU; they added that developing specific and valid criteria for underserved areas would supercede the IMU approach, which is based on a consensus among experts that has no explicit empirical basis They also suggested that national health interview survey data could be incorporated into measures used to predict the proportion of the area's population that was underserved (notwithstanding that this survey underestimates utilization of health services).

More recently, the General Accounting Office (GAO) has severely criticized the IMU in both construct and application. The GAO report, noting the Kleinman and Wilson assessment, also criticized the following points: the arbitrary use of a median IMU score to designate MUAs is likely to result in areas being designated as MUAs that have adequate primary medical care services; rational service areas are not considered in the development of MUAs; inappropriate data applications occur when county-wide rates are applied to census tract levels without adjustments; and finally, MUA designations have not been updated to reflect current data sources. The GAO report criticized grant applicants for similar deficiencies, as well as the grant application review process. In addition to recommending improvements in grant review, GAO suggested that the need for health programs and services be addressed from two perspectives: From the national perspective, need should be defined and mechanisms established to screen out geographic areas that meet the essential conditions of Federal support; from the state or local perspective, the indications of need resulting from the MUA or similar mechanisms should be modified with analyses that "realistically identify the need and expected demand for services in a given geographic area."[33] GAO did not, of course, offer precise guidelines for implementing this difficult recommendation.

In October 1981, the Bureau of Community Health Services (BCHS) proposed in the FEDERAL REGISTER to delete a number of previously listed medically underserved areas. The proposed deletions would result from recalculations of IMUs based on more recent data, and, more significantly, on adjustments to the weights applied to the actual values of two of the indicators—infant mortality and primary care physicians per 1000 popula-

[32] Kleinman, Joel C. and Ronald W. Wilson. *Are 'Medically Underserved Areas' Medically Underserved?* HEALTH SERVICES RESEARCH, 12:147-162, Summer 1977.

[33] Comptroller General of the United States, 1981, *op. cit.*

tion—"in order to reflect the changes in those indicators nationally since 1975."[34] About 560 counties nationwide would have been deleted from the original list; about 90 community health centers would lose Federal funding as a result of the elimination of MUAs.

Contacts with state health agencies and with BCHS have indicated that states and localities have responded to the proposed deletions in a variety of ways. Many states wrote preliminary letters indicating that they were reviewing the situation in their states and would soon provide more detailed responses. Some responses asked for specific proposed deletions to be kept on the list. Several states have analyzed in-depth their own state needs. The general state thrust has been that the data used by the BCHS are out of date; in some instances, the states have supplied their own more recent data, such as 1980 census counts or special studies they have conducted. The final result is expected to be that those MUAs on which no comments are received will be removed from the list of MUAs, but that a number of other MUAs will be restored as a consequence of new information.[35]

The Health Manpower Shortage Area (HMSA). The HMSA designation has been the principal measure for targeting Federal resources to redistribute health manpower. HMSA criteria have been established for several categories of health manpower, including primary care physicians, dentists, psychiatrists, veterinarians, and podiatrists. This discussion will focus on only the primary care criteria, since these are most relevant to the issue of medical underservice. Section 332(A) of the Public Health Service Act defined the term "health manpower shortage area" as "(A) an area in an urban or rural area (which need not conform to the geographic boundaries of a political subdivision and which is a rational area for the delivery of health services) which the Secretary determines has a health manpower shortage, (B) a population group which the Secretary determines has such a shortage, or (C) a public or nonprofit private medical facility or other public facility which the Secretary determines has such a shortage."

To be designated as an HMSA under the 1978 Federal regulations, a geographic area must satisfy three criteria: (1) The area must be a "rational service area" for the delivery of primary health care; (2) it must have a population:primary care physican ratio of at least 3500:1, or a population:physician ratio less than 3500:1 but greater than 3000:1, and either unusually high needs for primary medical services or insufficient capacity of existing primary care providers; and (3) primary care

[34] Federal Register 46(198):50677 (Oct. 14, 1981).
[35] Personal communication, Andy Jordan, Bureau of Community Health Services, Jan. 5, 1982.

medical manpower in contiguous areas must be demonstrated to be overutilized, excessively distant or inaccessible to the population of the area under consideration.

A population group may be designated as an HMSA for primary medical care if it can be demonstrated that access barriers prevent its members from using providers in the geographic area in which they live. The population group criteria are met automatically if the group consists of American Indians or Alaska natives who are members of Indian tribes. Facilities eligible to be designated as HMSAs include medium- and maximum-security Federal and state correctional institutions and public or non-profit private facilities. Such facilities must have a minimum number of inmates, or a minimum average daily inpatient census, as well as a critical value of full-time equivalent manpower.

The critical ratio for primary care shortage designation of 3500:1 was chosen because it is 1.5 times the mean population to primary care physician ratio by county for 1974 and because it selects approximately the lowest quarter of the country by county. The relative adequacy ratio, 3000:1, which is the minimum indicator of shortage for areas with unusually high need or insufficient capacity of existing providers, was chosen because it was the value midway between the shortage ratio of 3500:1 and the approximate median value for the country in 1974 of 2500:1. That both ratios can be explained does not diminish their arbitrary nature. For purposes of calculating the physician:population ratio, the Federal regulations specify rules for adjusting both numerator and denominator. Rules for converting various types of physicians to full-time equivalent counts are given, including adjustments for interns, residents, foreign medical graduates, part-time physicians, or those with restricted practices. In addition, consideration of the contribution of nurse practitioners and physician assistants can be considered where data are available. Rules for adjusting the population denominator of primary care physician:population ratios take account of different utilization rates by twelve different age-sex groups within the population; they also use a weighted sum to account for transient populations, including seasonal tourists and migrant workers.

Where an area's population:manpower ratio falls between the critical ratio of 3500:1 and subcritical ratio of 3000:1, a demonstration of unusually high need or insufficient capacity requires that the area meet at least one need criterion or at least two insufficient capacity criteria. The need criteria are: high birth rate (overall or teenage); high infant mortality rates; and a high percentage of the population in poverty. The insufficient capacity criteria include high physician visit rates, long waits for appointments (or waiting time in office or clinic), excessive use of emergency facilities for routine care, rate of physicians

not accepting new patients, and a low annual office visit rate. Finally, the regulations specify a ranking system for degree of shortage, using four groupings. Group 1, with the highest degree of shortages, includes areas with no primary care physicians or a ratio greater than 5000:1 physician with a demonstration of high need or insufficient capacity. Group 4, with the lowest degree of shortage, includes areas with 3500-4000:1 primary care physician or a population between 3000 and 3500 with the demonstration of need or insufficient capacity. Similar rankings are used for classifying facilities by degree of shortage.

Implementing the designation process for HMSAs involves not only the Department of Health and Human Services (DHHS), but also state and local review, using recommendations by health systems agencies, state health planning and development agencies, and the governor. In addition, any individual or agency may recommend to DHHS the "underserved" designation of a geographic area, a population group, or a facility. This recommendation would go through the same process of review and comment at the Federal, state, and local levels.

An in-depth assessment and critique of the HMSA criteria by researchers at Mathematica Policy Research (MPR) cited as crucial the failure to provide a narrative definition of a health manpower shortage area, particularly with respect to its characteristic market conditions.[36] This criticism is understandable in light of the existing rather complex HMSA regulations, which have been simplified here. The MPR reseachers concluded that the HMSA criteria try to satisfy several contradictory goals, thus no simple definition of health manpower shortage area is possible. The policy origins of the HMSA definition lie in the complex legislative history of manpower programs, which also contains inconsistencies.

The population:manpower ratio, the central component of the HMSA criteria, chiefly addresses availability of manpower. The focus is on identification of areas where availability, not accessibility, is the problem.[37] Nevertheless, the Health Professions Education Assistance Act of 1976 (P.L. 94-484) did add language to the HMSA criteria authorizing the use of indicators of need for health services by residents of the area (or population group or facility) being considered for designation, with special attention to indicators of infant mortality, access to health services, and health status. Clearly, this language goes beyond the previous focus on availability.

The Mathematica Group concluded that the application of the regulations could result in designation of three different types of shortage areas: economic shortage areas, normative

[36] Evaluation of Health Manpower Shortage Criteria, 1980, *op. cit.*
[37] Stambler, Howard V. In Proceedings of the Workshop on Health Manpower Shortage Areas, *op. cit.*, 55-59.

shortage areas, and nonshortage areas. An economic shortage area would be one with a low manpower availability and excess effective demand for health care services; a normative area would be one with high unmet needs, but an insufficient economic base to translate such needs into effective demand. Finally, a nonshortage area could be designated under the current criteria on the basis of low manpower availability, even though neither high need nor insufficient capacity existed. This seemingly anomalous result could occur if the population were relatively healthy or if the health manpower in question were particularly productive. For example, high rate of utilization of physician extenders might achieve such a result.

Structurally Underserved Populations. The Robert Wood Johnson Foundation has in recent years supported programs and efforts that attempt to improve people's access to general medical services, in rural underserved areas as well as in underserved inner cities. In addition, the Foundation, jointly with the U.S. Department of Health and Human Services, funded a major national survey of access to medical care that suggested that 12-15 million individuals have serious problems in obtaining needed medical care. The Foundation president, David E. Rogers, has used the term "structurally underserved" to refer to these populations.[38] Some specific demographic characteristics associated with lack of access to medical care include race, ethnicity, poverty, old age, and residence (inner city as well as rural). Dr. Rogers stated that the greatly increased supply of physicians expected in the U.S. over the next decade will not necessarily make personal physician care available to the structurally underserved populations. He cited four types of barriers to access for these groups. First, geographic isolation, either in the "backwaters of rural America" or in the seriously blighted inner-city areas, where inadequate public transportation, crime, and lack of basic public services interact to create special problems of access. Second, cultural isolation keeps many outside the traditional mainstream of American medicine. Third, poverty remains a barrier despite the Medicaid program (current and future cutbacks in Medicaid funding levels will probably exacerbate this problem). Finally, the structurally underserved include persons outside the mainstream because of frequent moving, e.g., migrant laborers, or because of isolation in mental institutions, homes for the aged, prisons, or other institutions.

The Foundation has no firm criteria for deciding whether an area is underserved, and therefore an eligible recipient of Foundation funds. It does use established criteria such as MUA and HMSA, but it also looks beyond these criteria. In the last

[38] Rogers, David E. Statement prepared for presentation to the Subcommittee on Health and Scientific Research, Committee on Labor and Human Resources, U.S. Senate, Sept. 24, 1980.

few years, the Foundation has tended to target specific populations and to support programs that will serve them. The Foundation staff still view existing measures as inadequate in assessing need and effective demand.

State Approaches to Defining Medical Underservice

State agencies have been as inconsistent as their Federal counterparts in defining underservice; a variety of methodologies have also emerged from state organizations to accomplish the task. Some merely modify the Federal criteria; others attempt to respond to local policy issues and problems. A recent Health Resources Administration study discerns different patterns in state and Federal medically underserved area designation processes.[39] Some state methodologies attempt a more rational service area designation, a concept of economic viability (some form of effective demand), and incorporation of other population characteristics based on specific needs or program applications of their designation processes. Their criteria identify a greater number of shortage areas than the Federal process.

In a recent study of 145 non-Federal programs to redistribute physician resources, 53 programs in 32 states were identified that employ a measure of medical underservice in their program administration.[40] The use of an underserved designation is mandated by statute in 23 of the programs. This subset generally provides individual or institutional financial incentives for medical practice or facility development in these designated areas. Four other health planning programs were identified that also use an underservice designation in their administration. Of these 57 programs in 34 states, only 8 rely strictly on HMSA and MUA criteria; roughly one-quarter of the 49 programs that developed their own methodologies use the Federal criteria as a starting point in their designation process, or Federal HMSA designations may *de facto* meet the state's criteria. (See Table 1.) Excluded from this data set are those programs that acknowledge a Federal designation of underservice (usually HMSA) as a convenient administrative tool. Undergraduate medical education preceptorship programs, which try to expose their students to medical care in underserved areas, are an example.

State-developed methodologies to identify underserved areas can be classified according to the concepts programs

[39] Personal communication with Marian Fox, Bureau of Health Manpower, Dec. 1981.

[40] Carpenter, E. and D. Seidman. National Compendium of Non-federal Physician Distribution Programs. Statewide House Officers Training System, Inc. Lansing, Michigan, Dec. 1981.

TABLE 1: Programs Employing Federal and State Underservice Methodologies*

A. Federal HMSA/MUA Designations (8 programs)		
Arkansas	Florida	New York
Colorado	Kentucky	Pennsylvania
B. State Methodologies (49 programs)**		
Alabama	Maryland	North Dakota
Arizona	Massachusetts	Ohio
Arkansas	Minnesota	Oklahoma
California	Mississippi	Oregon
Georgia	Missouri	Pennsylvania
Idaho	Nebraska	South Carolina
Illinois	Nevada	South Dakota
Indiana	New Jersey	Texas
Kansas	New Mexico	Utah
Kentucky	New York	Virginia
Maine	North Carolina	Wisconsin

* Some states had more than one program.
** About 25% of these methodologies are based on the Federal HMSA or MUA criteria.

incorporate and the measures employed to judge severity. (See Tables 2 and 3.) Fourteen states have developed concepts of rational service areas; for several reasons, geo-political boundaries remain the most popular mode of identifying underserved areas. Data on physician supply, population size, and demographic characteristics are collected according to geo-political boundaries, most notably by county. Many states cite the lack of useful information to define medical trade areas as an obstacle to constructing rational service areas. For this reason, the state of the art is limited to grouping geo-political areas according to identified patterns of medical care seeking. One state considers "established neighborhoods and communities which display a strong self-identity (as indicated by a homogeneous socioeconomic or demographic structure and/or a tradition of interaction or intradependency), have limited interaction with contiguous areas, and which, in general, have a minimum population of 20,000."[41] Isolated rural areas, deemed unable to finance a spectrum of health care services and providers, have been grouped into clinic service areas on the basis of travel time to the nearest secondary care facility. Several programs account for availability of services in contiguous areas in measuring underservice in a given rational service area.

[41] New Jersey Physician Dentist Loan Redemption Program Regulations, Section 9:16-1.6.

Five states have incorporated economic viability or effective demand into the designation process. A minimum population size, or number of patient visits necessary to support a provider or range of providers is specified, and areas are designated on the basis of this criterion.

Six states also have employed an urban/rural distinction to classify underserved areas, applying different criteria to measure underservice in each setting. For example, a Missouri program bases urban underservice on a physician:population ratio, while rural underservice is identified by maximum population size.

Some states also define underservice as shortages of various medical specialties and as institutional settings with chronic personnel shortages. Nine states allow medical graduates who have received financial aid on the condition that they will practice in underserved areas to fulfill their commitment through in-state practice in family medicine, primary care, and/

TABLE 2: Underservice Concepts in State Programs*

A. Rational Service Area (15 programs)		
California	Nebraska	Pennsylvania
Idaho	New Jersey	South Carolina
Kansas	New York	Texas
Maryland	Ohio	Wisconsin
Minnesota	Oklahoma	
B. Economic Viability (5 programs)		
Idaho	New Jersey	Wisconsin
Kansas	North Carolina	
C. Urban/Rural (7 programs)		
California	Florida	Indiana
Colorado	Illinois	Missouri
D. Specialty (10 programs)		
Alabama	Kansas	New York
California	Kentucky	Texas
Georgia	Nebraska	Wisconsin
E. Institutions (8 programs)		
Alabama	New Jersey	Oklahoma
Georgia	New York	
Illinois	North Carolina	

* Some states had more than one program.

TABLE 3: Underservice Measures in State Programs*

A. Physician:Population Ratio (37 programs)		
Arizona	Indiana	North Carolina
Arkansas	Kentucky	Ohio
California	Maine	Oklahoma
Colorado	Massachusetts	Oregon
Florida	Nebraska	Pennsylvania
Georgia	Nevada	South Carolina
Idaho	New Jersey	South Dakota
Illinois	New Mexico	Virginia
1. With Numerator Modifications (7 programs)		
Indiana	Kentucky	New York
Kansas	Minnesota	Wisconsin
2. With Denominator Modifications (1 program)		
Indiana		
B. Population Size (12 programs)		
Alabama	Kansas	North Carolina
Arkansas	Minnesota	North Dakota
Georgia	Missouri	Wisconsin
Illinois	New Mexico	
C. Direct Measures of Need (17 programs)		
Alabama	Massachusetts	Oklahoma
California	Nebraska	Oregon
Florida	Nevada	Pennsylvania
Indiana	New York	Utah
Maine	North Carolina	
Maryland	Ohio	
D. Minimum of Personnel (1 program)		
South Carolina		

* Some states had more than one program.

or public health, with or without geographic restriction. Wisconsin applied the statistical technique of cluster analysis to the distribution of specialists in acute care service areas to identify those areas that may lack a given type of specialist service. State or non-profit health facilities experiencing chronic vacancies, and mental health and corrections institutions serve as acceptable sites for fulfillment of practice commitment of 8 programs in seven states.

Of the 57 programs studied, 45 use some form of physician:population ratio to determine medical underservice. Thirty-seven programs use a simple count of the number of physicians per population in constructing the ratio, including those programs that strictly adhere to the Federal HMSA and MUA criteria. An additional 7 programs modify the numerator of the ratio by estimating full-time equivalents of physicians, identifying number of patient visits available in the physician pool, excluding from the count physicians over age 65, or giving different weights to particular specialties to construct a primary care physician:population ratio.

Methods for calculating full-time equivalents vary, but generally include a weighting factor for percent of time spent on direct patient care. Full-time equivalents are generally calculated for each specialty group. Information on time spent with patients and in specialty areas is provided by the physicians themselves, typically through a questionnaire in conjunction with licensure or relicensure forms. Criteria for determining necessary number of physician visits required by a population are acquired either by local studies (Maryland) or from the results of the Health Interview Survey of the National Center for Health Statistics (NCHS). An additional criterion used by one program considers an area underserved if it lacks a minimum number of personnel, in this case at least one primary care doctor. One program modified the denominator of the ratio, the population count, by weighting for the age and income level of the population.

The remaining 12 programs use a maximum population size as the criterion for identifying areas of underservice. These programs focus on rural areas and assume that an area below a "critical" population level lacks necessary medical services.

By examining economic, social, and demographic characteristics of the population, 17 programs estimate need for health services by proxy measures in addition to using population size and physician:population ratios for identifying underserved areas. Programs that use the Federal MUA criteria are included in this subset. The indirect measures of health status include age structure (percent young, percent elderly), income level (percent of population on welfare/below the poverty level), race/ethnicity factors (percent non-English-speaking) and fertility data (percent of women ages 15-44, percent births to teenage women). Mortality data is exclusively incorporated in the form of infant mortality rates. Health status measures are never the only criteria used in defining areas of underservice.

Few states have taken the initiative in the development of methodologies to identify medically underserved areas, and designating areas in need of health care resources is a requirement for few state health programs. Further, the structure of the Federal project grant programs, a major subset of programs that

use the HMSA/MUA designations for allocating resources, places state governments in a reactive role of responding to Federal directives. Only two programs using the Federal designations of underservice allocate funds directly to the states; they represent a meager portion of the total funding available to those programs requiring an underservice designation. State and local health planning agencies gathered data and supporting materials in the HMSA and MUA designation processes in order to garner limited health resources from Bureau of Community Health Services project grants and the National Health Service Corps. In the present structure, state agencies have had to struggle even to win a coordinating role in programming of Federal project grants.[42] The reactive position of state planning agencies is reinforced by the fragmentation of responsibility within the state for designating underserved areas. Local HSAs are charged with data collection responsibilities, and in large, populous states, competing agendas of several HSAs are an additional barrier to the development of coherent state health policy on medical underservice.

The private sector, represented chiefly by state and local medical and dental societies, has also participated in designating medically underserved areas. The process of designating critical health manpower shortage areas for the placement of National Health Service Corps personnel directed comprehensive health planning agencies to identify those areas meeting the state criteria. However, the Emergency Health Personnel Act Amendments of 1972 required certification by the appropriate health professional society as to the need for assignment of Corps personnel. The 1976 Health Professions Educational Assistance Act, in revising the assignment of Corps personnel, makes no mention of this certification procedure. In fact, health professional societies are not explicitly referred to in any portion of the legislation dealing with medically underserved area designation.

Several developments may have led to the decreased weight placed on private sector input in the designation process. The greater emphasis on the role of state and local government units in health planning, as well as other structural changes effected by the 1974 National Health Planning and Resources Development Act, have strengthened the voice of consumers in local and statewide health planning.[43] In the early years of the Corps, medical and dental societies may have been the most valuable and reliable information sources on supply and distribution of health personnel. Increased data availability

[42] Madison, Donald. *A View of Rural Health Services.* Am. J. Public Health, 71:89-90, Jan. 1981 (Supplement).

[43] Altman, Drew. *The Politics of Health Care Regulation: The Case of the National Health Planning and Resources Development Act.* J. Health Politics, Policy and Law, 2(4):560-580, Winter 1978.

fostered by the Cooperative Health Statistics System and the infusion of resources, money, and professionalism into Health Systems Agencies (HSAs) created an alternative authoritative source for assessing local needs.

The participation of professional societies in addressing issues of medical underservice must not be judged solely upon their involvement in the Federal process or upon examination of statutory requirements of state programs similarly involved. A picture can be completed only by a case-by-case examination of state programs concerned with the availability of health resources in underserved areas. The 1981 study of physician distribution programs found several examples of a high degree of involvement of state medical societies, specifically in loan programs to medical students that forgive portions of the principal and interest in return for practice in underserved areas. The role of the private sector has changed but not disappeared since the early stages of the National Health Service Corps. In specific instances, state or local medical societies continue to protest placement of Corps personnel. Through their representation on HSA boards and SHCCs and in direct interaction with state legislatures, professional associations also influence policymakers' perceptions of the problem of medical underservice as well as governmental efforts to solve it.[44]

Appropriateness of Existing And Proposed Indices of Underservice

Appropriateness of Standards and Issues of Equity. The number of people who reside in medically underserved areas varies widely depending on the index or standard applied. Current HMSA criteria designated 1921 primary care shortage areas, including populations and facilities, which contain almost 42 million people. On the other hand, the Robert Wood Johnson Foundation, relying on its national survey data on access to medical care, has estimated that about 12-15 million people are "structurally underserved." State methodologies tend to include more areas within the underserved or health manpower shortage rubric than the Federal criteria. California, for example, has designated 159 rural and 64 urban critical health manpower shortage areas (CHMSA), which contain over 6 million people. In contrast, the 32 Federal HMSAs designated in California have a population of 860,000.

Current methods may designate as underserved some adequately endowed areas (false positives) and may omit areas and populations experiencing actual underservice (false negatives). Such errors of designation have important implications for the

[44] See, for example, R.V. Crawford, *State Health Care Policymaking: The Tennessee Primary Care Act of 1973*. J. Health Politics, Policy and Law, 4(4):691-702, Winter 1980.

equitable distribution of health resources. A basic criterion for redistributive programs is target efficiency, which includes two dimensions: vertical and horizontal. Vertical efficiency denotes the ability of a program to provide benefits only to the target group, in this instance, medically underserved areas and populations. Horizontal efficiency denotes the ability of the program to reach all of the class of individuals to whom the redistributive effort is directed; e.g., to reach all the medically underserved areas or populations. The concept of vertical efficiency relates to the inequities arising where benefits are redistributed to non-needy individuals; horizontal efficiency is related to equity in terms of equal treatment of equals. Horizontal efficiency also relates to the norm of adequacy, since it can be defined in terms of the total amount of benefits "needed" by that group.[45]

As Federal funding diminishes for health and social welfare programs, efficient allocation of resources becomes a critical component of the concept of equity in health. The lack of consensus about appropriate indicators of medical underservice implies that the criteria in use at the Federal and state levels discriminate poorly among target populations for purposes of either vertical or horizontal efficiency. Therefore, they have serious implications for norms of equity as well. The failure to gain agreement on a standardized definition of medical underservice stems, not only from the fuzziness of the meaning of such terms as "access," but also from the ambiguity and conflict embodied in the goals and objectives of many of the programs using underservice designation as a decisional tool. To these normative concerns must be added the technical and methodological issues involved in the measurement of criteria for underservice.

Goal Ambiguity. The fact that programs employing indices of medical underservice often have vaguely specified or multiple goals is reflected in the designation criteria themselves. The pervasiveness of physician:population ratios in Federal and state criteria might suggest that availability, alone, is the central goal of redistributive efforts. But the legislative and administrative histories of these programs clearly show a hope and expectation that redistributing physician resources will help eliminate disparities in utilization of health services and ultimately improve health status.[46] Until what is meant by underservice is clearly specified, it is impossible to judge the appropriateness

[45] Weisbrod, B.A. *Collective Action and the Distribution of Income: A Conceptual Approach.* In Haveman, R.H. and J. Margolis, PUBLIC EXPENDITURES AND POLICY ANALYSIS. Chicago: Markham Publishing Company, 1970, 124-128.

[46] See, for example, the discussion of the formulation of the National Health Service Corps legislation in Eric Redman, THE DANCE OF LEGISLATION. New York: Simon and Schuster, 1973.

of criteria for designating underserved areas. As Hadley has noted, availability measures are widely used, not because they are appropriate to the objectives of greatest interest, but because they are readily obtained from existing data.[47]

The earliest conceptualization of underservice saw it in terms of spatial and geographic factors, and as a predominantly rural problem. In the case of state programs of medical scholarships and loans with service payback conditions, underservice was sometimes defined simply in terms of size of community; e.g., population of 10,000 or less. A growing recognition of the multiple causes of medical underservice has led to the evolving complexity of criteria to measure it. Medical underservice in urban areas is now acknowledged to affect a large number of people than rural underservice and to result more from economic, social, and cultural barriers than from spatial factors. Moreover, the most critical problems of underservice result from the interaction of spatial and socioeconomic barriers to access as exemplified by the rural poor, inner-city minority populations, and migrant workers.

For effective policymaking, designation criteria need to identify the multiple causes of maldistribution of resources and the barriers to access, and the criteria must incorporate these causes and barriers in standards. In fact, policy instruments have been chosen before the standards for measuring underservice were developed, or criteria developed for one program have been adopted for use in other programs, which may address different components of underservice.

Technical and Methodological Issues. A number of technical issues are potential obstacles to the application of any index of medical underservice, including data availability, practical methodologies, and validity and reliability of indicators. Either in devising their own criteria for medical underservice or in responding to the Federal criteria, state health planners complain that the data are obsolete or inaccessible. Moreover, the application of shortage area criteria to small geographic areas presents problems of reliability when current data are lacking. A shift in location of a single physician in a small geographic area can move that area in or out of a shortage designation. Criteria that require the collection of special survey data or other extensive data acquisition, such as patient origin studies, are too costly for many state and local programs. Data elements must also be available for actual medical service areas. In most cases this will require disaggregation of the data below the county level.

Although the literature describes a number of complex econometric models for assessing need or demand for health

[47] Hadley, Jack: *Alternative Methods of Evaluating Health Manpower Distribution*. MED. CARE 17(10):1054-1060, Oct. 1979.

services, such models are absent from the underservice criteria currently employed. In most cases, the technical capabilities for their application are simply not widely enough available to make these methodologies useful in program administration. They may also demand data capabilities that are not developed or are too expensive to be practical for state and local administration. Because the application of underservice criteria decide the allocation of public resources, it is particularly important that methodologies are understandable and have certain face validity. Where the methodology seems so complex that it cannot be adequately monitored by program administrators or so loose that it can easily be circumvented (a not uncommon perception of HMSA criteria at the state level), the political consensus necessary for its acceptability will be lacking.

The question of the validity of criteria for underservice essentially relates to whether the criteria measure the concept. To risk repetition, the concepts being measured are not clearly delineated in either of the current Federal underservice indices, nor is there evidence that any more clarity exists at the state level. If we take the common-sense approach that medical underservice designations are intended to direct the allocation of resources toward programs that improve the people's health, the area in which validity is most problematic relates to measures of health status. Vital statistics data, in particular, mortality data, are probably the highest-quality readily available data on health status outcomes (as measured by infant mortality rates, premature deaths, etc.). Nevertheless, there are serious concerns about their usefulness in directing allocation of resources to health care services. The marginal contribution of medical care to life expectancy and other mortality-based measures of health status has been noted by Fuchs, among others.[48] At best, death rates are an inadequate indication of medical care outcomes.

Other deficiencies of mortality data are that they tell us nothing of differences in function or quality of life.[49] Function may improve as a consequence of the ability to obtain medical care, which is clearly a matter of interest in achieving the goal of equity in health.

Utilization rates are widely used to measure both demand and need for medical care, but these rates are highly sensitive to variations in practice patterns in different geographic areas, as well as to differences in organizational and payment modes. Similarly, professional assessments of medical service requirements, as exemplified by the GMENAC methodology, have the potential for error; these errors arise from the uncertain rela-

[48] Fuchs, Victor R. *Some Economic Aspects of Mortality in Developed Countries.* In Economics of Health and Medical Care, 174-193.
[49] Mushkin, S.J. and D.P. Wagner. *Expected Mortality as a Criterion in Health Policy Evaluation.* Public Services Laboratory, Report No. B12, Georgetown University, Nov. 1977.

tionship between medical interventions and health care outcomes. This uncertainty is reflected in the wide variety of professional norms of treatment that currently exist. Indeed, as Wennberg and Gittelsohn have pointed out, even with less uncertainty about outcomes, physicians and patients might still value differently the outcomes of alternative treatment modalities.[50]

Finally, since the designation of medically underserved areas necessarily must focus on relatively small geographic units, the potential for measurement error is increased. Particularly where criteria rely on incidence of medical conditions or mortality causes of relatively low frequency, reliability will be low unless moving averages or multiple-year averages or similar norming techniques are used.

Similar measurement issues arise in considering the criteria for determining resource availability. Reasonable, consistent, and uniformly applicable rules for converting manpower resources into standardized units of medical service output are needed, including adjustments for use of nonphysician personnel, and for average hours worked and other indices of productivity.

Physician Location Decisions As An Element of Medical Underservice

A good deal of research in physician location decisions has been conducted over the past two decades, due at least in part to the perception that medical underservice is largely a problem of physician maldistribution. Certainly, understanding the factors that influence physician distribution is important in the formulation of policies to address underservice, if those factors are appropriate objects of policy intervention. For example, much data on the social and psychological determinants of physician career choices may be of limited usefulness to policymakers.

A brief overview of the literature on physician maldistribution highlights those research findings that have influenced the formulation of public policy. A useful classification of causal factors is provided by Rushing: individual characteristics of physicians; supply of physician manpower and the demand for medical care; location (and characteristics) of training institutions; medical characteristics of communities; and organizational aspects of medicine.[51]

---

[50] Wennberg and Gittelsohn, 1980, *op. cit.*, citing McNeil, F.J., *et al. Fallacy of the Five-Year Survival in Lung Cancer.* NEW ENG. J. MED. 299:1397-1401, Dec. 1978.

[51] Rushing, W.A. COMMUNITY, PHYSICIANS, AND INEQUALITY. Lexington MA: Lexington Books, 1975.

Preferences formed before attending medical school, based on family and community background and socialization, may be either strengthened or altered during medical training. On the one hand, physicians tend to practice in a community similar in size to the one reared in, particularly with respect to rural practice.[52] On the other hand, a number of studies have reported that most medical schools socialize students to practice in urban areas.[53] Selective recruitment of medical students from rural and minority backgrounds has been suggested as a means of increasing the supply of physicians willing to practice in underserved locations.[54] The 1971 Health Manpower Training Act and the more recent GMENAC study both enunciated this goal.

In the absence of reinforcing socialization during medical training, selective recruitment is not likely to achieve the desired results. Choice of specialty interacts with choice of practice location, and the former is heavily influenced by medical faculty role models.[55] In particular, the choice of family medicine as a specialty is complementary to selection of a rural practice site, underscoring the need for faculty role models and training experiences in rural sites.

In light of research findings that physicians prefer to treat patients from backgrounds similar to their own, the current underrepresentation of minorities among medical school entrants is a source of concern. Longitudinal studies of minority physicians in California have persuaded health manpower analysts in that state that increased medical school enrollment of racial and ethnic minorities is the most effective way to improve availability of medical services to underserved minority populations.[56] A number of states have proposed cutbacks in

---

[52] Bible, B.L. *Physicians' Views of Medical Practice in Non-metropolitan Communities.* Public Health Rep. 85:11-17, Jan. 1970. Cooper, J.K., K. Heald, and M. Samuels. *Affecting the Supply of Rural Physicians.* Am. J. Public Health 57:756-759, Aug. 1977.

[53] Breisch, W.F. *Impact of Medical School Characteristics on Location of Physician Practice.* J. Med. Ed. 45:1068-1070, Dec. 1970. Bynum, G., R. Sanchez, and W. Odegaard. *Medical Education: A Causal Agent in Physician's Maldistribution.* J. Med. Ed. 47:722, Nov. 1972.

[54] Cullison, S., C. Reid, and J.M. Colwill. *Medical School Admissions, Specialty Selection, and Distribution of Physicians.* J.A.M.A. 235:502-55, Feb. 2, 1976.

[55] Bynum, *et al.*, 1972, *op. cit.* Institute of Medicine, A Manpower Policy for Primary Health Care. Washington, D.C.: National Academy of Science, May 1978. Taylor, M., *et al. Medical Students' Attitudes Toward Rual Practice.* J. Med. Ed. 48:885-895, Oct. 1973.

[56] For example: Mayfield, Sharon and Edwin P. Merwin. Black Physicians in California. Office of Statewide Health Planning and Development, Los Angeles County, 1980 Practice Location Correlated to Crit. Health Manpower Shortage. April 1981. Faustino, Bernadette, Edwin P. Merwin, and Frank Meza. California Chicano Physicians. 1967-1977 Future Plan for Medical Specialty, Practice Location and Patient

support for medical education, in the face of constrained revenues and forecasts of physician surpluses by 1990.[57] These cutbacks could adversely affect recruitment of minorities into medical careers, unless specific steps are taken to protect and improve on the gains made since the 1960s.

Restructuring undergraduate and graduate medical education is another approach to altering physician career choices. Examples of educational innovations include area health education centers, rural preceptorships, decentralized training sites, and required course work in family medicine. Participation in preceptorships, for example, has been found to be associated with choice of rural practice, but it is not clear whether prior preferences of participants were simply reinforced by the experience.[58] The growth of family medicine departments has also been seen as important in altering the specialty mix of physicians and improving physician distribution.[59] Perhaps the strongest case can be made for recruiting students with preferences for rural practice and family medicine and reinforcing these preferences with appropriate role models and educational experiences.[60]

Weiskotten's longitudinal studies of physician careers have emphasized the importance of location of graduate training as a determinant of place of practice.[61] However, others have argued that preferences for location of practice are formed prior to graduate training, and that these preferences determine the site both of graduate training and eventual practice.[62] For much

---

PROFILE. Office of Statewide Health Planning and Development, April 1981. Becerra, Maria Elena, Rosana Madrid, and Edwin P. Merwin. *California Chicano Physicians 1969-1977 Actual Choice of Medical Specialty, Practice Location and Patient Profile*. Office of Statewide Health Planning and Development, April 1981.

[57] See, for example, Office of Health and Medical Affairs. PLANNING FOR PHYSICIAN REQUIREMENTS AND SUPPLY IN MICHIGAN. Michigan Department of Management and Budget, Lansing, MI, May 1981.

[58] Steinwald, B. and C. Steinwald. *The Effect of Preceptorship and Rural Training Programs on Physicians' Practice Location Decisions*. MED. CARE 13:219-229, March 1975.

[59] Mayo, F., *et al. Graduate Follow-up in the Medical College of Virginia/Virginia Commonwealth University Family Practice Residency System*. J. FAM. PRACTICE 11(5):731-742, 1980.

[60] Verby, J.E., *et al. Learning Forestry out of the Lumberyard: A Training Alternative for Primary Care*. J.A.M.A. 246(6):645-47, Aug. 7, 1981. Boulger, J.G. *Family Practice in the Predoctoral Curriculum: A Model for Success*. J. FAM. PRACTICE 10(3):453-458, 1980.

[61] Weiskotten, H.G., *et al. Trends in Medical Practice: An Analysis of the Distribution and Characteristics of Medical College Graduates, 1915-1950*. J. MED. EDUC. 35:1071-1120, Dec. 1960. Reskin, B. and F.L. Campbell. *Physicians' Distribution Across Metropolitan Areas*. AM. J. SOCIOLOGY 79:981-98, Jan. 1979.

[62] Cooper, J.K., *et al. The Decision for Rural Practice*. J. MED. EDUC. 47:939-44, Dec. 1972. Lee, R.H. *Regional Physician Supply and Graduate Medical Education*. HEALTH POLICY AND EDUCATION, 1980.

graduate training in specialty areas other than primary care, relocation of graduate training may not be possible; the clinical facilities and patient mix required for residency programs dictate location in areas with appropriate facilities and population density. It may, however, be possible to move some training away from tertiary medical centers to community hospitals, as California is attempting to do under the Song-Brown Act.

Financial inducements in the form of scholarships and loans are another policy instrument widely used to influence physician distribution by canceling loan obligations for practice in underserved areas. Despite their prevalence, evaluations of most programs have found them ineffective.[63] As for economic incentives directed to practicing physicians (e.g., start-up loans, tax credits, or other subsidies), most programs are too small or too new to evaluate. Survey data indicate that physicians rank these incentives low in influence.[64]

The most significant economic factors associated with physician location are demand related: per capita income, health insurance coverage, per capita health care expenditures.[65] Hadley, reporting on eight econometric studies of physicians' location choices, found physicians tend to locate where physicians' incomes are highest.[66] These locations are also most richly endowed with clinical facilities, hospitals, ancillary services, continuing education opportunities, peer reinforcement, and referral networks, all of which are seen as essential to mainstream medicine as taught in most American medical schools. Thus, as Rushing has suggested, the medical charac-

---

[63] CONSAD Research Corporation. AN EVALUATION OF THE EFFECTIVENESS OF LOAN FORGIVENESS AS AN INCENTIVE FOR HEALTH PRACTITIONERS TO LOCATE IN MEDICALLY UNDERSERVED AREAS. A report prepared for the Department of Health, Education and Welfare, Office of the Secretary, Contract No. HEW-05-73-68, Washington, D.C., 1973. Mason, H.R. *Effectiveness of Student Aid Programs Tied to a Service Commitment.* J. MED. EDUC. 46:575-83, July 1971. Parker, R.C. and A.A. Sorensen. *Physician Attitudes Toward Rural Practice: Answers to Questions That Were not Asked.* FORUM ON MED.: 411-416, June 1979.

[64] Heald, K.A., *et al*, CHOICE OF LOCATION OF PRACTICE OF MEDICAL SCHOOL GRADUATES: ANALYSIS OF TWO SURVEYS. Rand Corp., Santa Monica, California, R-1477 Nov. 1974. Cantwell, J.R. *Implications of Reimbursement Policies for the Location of Physicians.* AGRIC ECON RES 31:25-35, April 1979. Fuchs, V.R. and M.J. Kramer. DETERMINANTS OF EXPENDITURES FOR PHYSICIANS' SERVICES IN THE U.S., 1948-68. DHEW Publication No. (HSM) 73-3013. Rockville, MD, Dec. 1972. Yett, D.E. and F.A. Sloan. *Migration Patterns of Recent Medical School Graduates.* INQUIRY, 11:125-142, June 1974.

[65] Coleman, S. PHYSICIAN DISTRIBUTION AND RURAL ACCESS TO MEDICAL SERVICES. (Executive summary). Prepared for DHEW, HRA, Division of Medicine under Contract No. 231-750-0613 by the Rand Corporation, Santa Monica, CA. No. R-1887/1, April 1976.

[66] Statement of Jack Hadley before the Committee on Labor and Human Resources, U.S. Senate, April 8, 1981.

teristics of communities and the organizational aspect of medicine interact to draw physicians to generously endowed areas, and to exacerbate the problems of poor, rural, inner-city, and economically declining areas.[67]

So long as medical underservice is perceived only in terms of physician maldistribution, policy initiatives will try to influence individual physicians' decisionmaking, a conspicuous example being financial incentives to foster the relocation of physician practices to so-called underserved areas. The literature on physician location choices and the past experience of these programs suggest pessimism about this approach to medical underservice. Replacing lenient buyout provisions with more stringent penalties for failure to perform required service, coupled with rising tuition levels at public medical schools, may increase the impact of financial aid programs. However, it can be argued that service-conditional aid treats the less affluent inequitably, since it is they who are most likely to need financial assistance and thus are more constrained in choice of practice location. Yet this is not an argument for abandoning financial support of low-income students.

Broader issues of social equity are raised by proposals such as that debated in Congress in 1974 to require of every medical graduate a period of service in designated shortage areas. The justification is that public subsidies to medical education are very high and that physicians as a class are one of the highest income groups in the country. Even allowing for their lengthy training, physicians experience a rate of return on their educational investment far higher than that of other occupational groups. Balanced against that argument is the severe restraint on individual freedom and autonomy represented by a service obligation imposed by governmental authority, in contrast to one voluntarily incurred. Responses to the proposed 1974 amendment to the Health Manpower Act and to state regulatory approaches to altering physician location patterns indicate a sense that such coercive measures are too high a cost to pay for more "equitable" distribution of physicians. Meanwhile, service-conditional financial aid programs continue to function at the state level, and several have modified their provisions to raise the cost and penalties of buying out of required service. In general, however, incorporation of service obligations into financial aid programs seems to be as much a means of gaining acceptability for large public subsidies for medical education as it is directed toward solving problems of maldistribution of resources. Equity issues aside, it is not clear that the present system of implicit and explicit subsidies and incentives are efficient in achieving a redistribution of physician resources. This is not to deny that the maldistribution of physician resources, in terms of both geographic distribution and the bal-

[67] Rushing, W.A., *op. cit.*

ance between primary care and other specialties, is an important obstacle to equity in access to health care. Rather, physician maldistribution may be more of a symptom than a cause of the problem of medical underservice.

As the nation faces an impending surplus of physicians, those who argue, like Schwartz *et al.*,[68] that market forces will redistribute physician resources to now underserved areas have gained support. It is probably safe to say that many currently designated HMSAs may soon be able to recruit and retain physicians. A number of privately funded recruitment and placement programs have been initiated in recent years, focusing in particular on rural locations.[69] Such approaches may accomplish a more efficient redistribution of physicians to "shortage" communities with sufficient economic demand to support a physician practice than have either the National Health Service Corps or other forms of obligated service.

There remain, however, backwaters of small, sparsely populated rural areas and the urban ghettos and barrios, which are not likely to benefit from changing physician market forces. Even the strongest proponents of the view that market forces will diffuse the supply of physicians recognize that some population groups will remain too impoverished or otherwise disadvantaged for market solutions. In the words of Schwartz *et al.*, "special intervention may be necessary on the part of government."[70]

In brief, it seems desirable to uncouple the development of rational criteria for designating underserved areas from the formulation of policies for the public support of medical education. Public support for medical education, in terms of its relationship to medically underserved populations, could then be responsive to the problem, rather than setting parameters for its definition. Solutions to the problem of medical underservice could then be proposed on the basis of elements of the problem that respond to policy intervention. Moreover, this approach would help distinguish between policies appropriate to governmental initiatives and those that might better be left to the private sector.

Conclusions

Overcoming medical underservice, in the sense of matching resources to needs, has so far eluded the efforts of health policymakers in the U.S. and other countries. The lack of agree-

[68] Schwartz, W.B., *et al. The Changing Distribution of Board-Certified Physicians.* New Eng. J. Med. 303:1032-1038, 1980.
[69] Carpenter and Seidman, *op. cit.*
[70] Statement of William B. Schwartz, Joseph P. Newhouse and Albert P. Williams to the Committee on Labor and Human Resources, U.S. Senate, April 8, 1981.

ment about what is meant by underservice and the difficulties of constructing valid measures of such concepts as need and demand reflect the fundamental uncertainties that underlie the relationship between medical care and health outcomes.

The rapid development and diffusion of new and expensive medical technology, often without rigorous assessment of its efficacy, heightens these uncertainties. Just as the proliferation of medical therapies makes defining "needs" more difficult, so it also intensifies the maldistribution of resources in two ways. Medical resources become more concentrated in populous, affluent urban areas able to support costly facilities and procedures. In turn, the disparity in resources allocated widens between primary care and prevention, on the one hand, and expensive therapeutic interventions on the other. The imbalance in the distribution of physicians by location and by specialty follows that pattern of resource allocation.

Thus, alleviating medical underservice may require addressing "overservice." This is not easy; although redistribution from overserved to underserved areas will increase social benefits attributable to reduced morbidity and mortality, it will increase private costs. Definitions of "underservice" in current use focus almost exclusively on deficiencies in health care resources compared with those available to the "average" area or population. In doing so, they direct attention away from the possibility that resources are misallocated to needless (and occasionally harmful) treatments.

An alternative focus sees underservice as a deficiency in availability of a basic floor of health care protection.[71] To incorporate that view into a definition of underservice, one would begin with observed differentials in health status outcomes among populations, using available measures of mortality, disability, acute and chronic illness, then relate those differentials to the availability and application of health care. This relationship can be estimated both empirically and with existing research findings and can be a guide to targeting scarce health resources to areas and populations by priority of needs. This approach seems preferable to relativist measures of resource distribution, which pursue and forever lag behind the proliferation of new technologies.

It would be naive to assert that better operational measures of underservice can assure a solution to maldistribution of health care resources. Nevertheless, the sense that present Federal criteria are unreliable and invalid is evident in state responses, recent Congressional deliberations, and several GAO

[71] For example, Norman Daniels, in *Equity of Access to Health Care: Some Conceptual and Ethical Issues*, Appendix B in Volume Two of this report, argues for a social obligation to provide a "basic tier" of health care services that meet important health needs as judged by their impact on people's range of opportunity.

reports. State methods of establishing the existence of medical underservice are even more various and often inconsistent with Federal criteria. Political support for redistributive policies depends on the belief that resources are being allocated fairly and efficiently. As guides to allocating scarce health care resources, indices of medical underservice now in use are inadequate because they fail to clearly state a standard for equity of acce

Allocation in Varied Medical Settings

IV

The Doctor, the Hospital, and the Definition of Proper Medical Practice

Anthony L. Komaroff, M.D.*

It has become commonplace to observe that the doctor largely determines the cost of hospital care. What is interesting is to speculate about how this fact can be used to advantage in an era of economic constraints. In this essay we will amplify on the commonplace and join the speculation. Mostly, however, we will discuss an inconvenient reality that makes it hard to pose, much less answer, the ethical questions raised in an era of cost containment: the difficulty of defining a patient's medical need and a doctor's proper medical practice. For in considering alternative schemes of reimbursement, new organizational arrangements between doctors and hospitals, or mechanisms for rationing, we cannot judge what is ethical without first knowing what is medically necessary.

The Doctor and the Hospital Administration

Of the $245 billion spent on health care in 1980, hospital care accounted for about 40% of the total.[1] Since the Second World War, a substantial fraction of the increase in medical costs (especially hospital costs) has resulted from an increased intensity of medical resource utilization per capita (in contrast to price inflation or to an increase in the number of patients

*Division of General Medicine and Primary Care, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; and the Center for the Analysis of Health Practices, Harvard School of Public Health, Boston, Massachusetts.

October 1981

[1] M Freeland, G Calat, CE Schendler, *Projections of national health expenditures,* HEALTH CARE FINANCING REV (1980, 1:1-27).

served).[2] In hospital care, the only significant trend towards reduced intensity of resource utilization is in reduction of the length of stay and hospital days, with the associated increased tendency to treat patients in an ambulatory setting. For example, Blue Cross and Blue Shield plans have witnessed a 19.9% reduction in the number of hospital days financed per thousand subscribers per year over the past ten years.[3]

The doctor plays a critical role in determining the size and the nature of expenditures for hospital care—deciding whether to hospitalize the patient and for how long, determining what diagnostic tests and treatments to order, and deciding whether to perform surgery. Thus, a hospital's medical staff is solely responsible for determining the utilization of all those goods and services for which a hospital can charge.

At the same time, the hospital administrators have an important role in determining the costs of hospital care. In particular, they make decisions that influence the cost of a day of care: determining the number of staff required to run the hospital (and their salaries), deciding whether to expand or renovate the physical plant, setting charges for all elements of hospital care, and deciding whether or not to obtain new diagnostic and treatment technologies for the institution.

The hospital trustees have final responsibility for the decisions of both the medical staff and the hospital administrators. As we shall discuss later, it is very difficult—both practically and politically—for the trustees to exercise the full measure of their authority.

The medical staff's critical role in attracting patients to the hospital and in making all of the decisions that affect the hospital's reimbursement gives it considerable influence on those decisions formally reserved to the hospital administrators and trustees. After all, members of the medical staff usually have the option of practicing at another hospital. Conversely, the hospital provides services that are essential for the doctor's practice. And the hospital—for reasons of prestige or convenience—may be in a position to compete favorably for the doctor's "business."

---

[2] *See:* HE Klarman, DP Rice, BS Cooper, *et al.*, *Sources of increase in selected medical care expenditures, 1929-1969*, DHEW Staff Paper No. 4, Washington, DC: Government Printing Office, 1970; A. Scitovsky, *Changes in the use of ancillary services for common illness*, in: SH Altman, R Blendon (eds.), MEDICAL TECHNOLOGY: THE CULPRIT BEHIND HEALTH CARE COSTS? DHEW Publication No. (PHS) 79-3216 (Washington, DC: Government Printing Office, 1979:39-56); RM Gibson, *National health expenditures, 1979*, HEALTH CARE FINANCING REV (1979, 1:1-36); AS Relman, *The allocation of medical resources by physicians*, J MED EDUC (1980, 55:99-104).

[3] WJ McNerney, *Control of health-care costs in the 1980's*, N ENGL J MED (1980, 303:1088-95).

Thus the medical staff, the hospital administration, and the trustees coexist in a state of harmonious or dissonant tension, as the case may be. In the 1980s, the tension is more likely to be dissonant. Hospitals are buffeted by two paradoxical forces. On one hand, a sluggish economy encourages constraints on the utilization of resources.[4] On the other hand, this era of belt tightening is also one of tremendous growth in the availability of new diagnostic and therapeutic technologies. Doctors and hospitals are expected to do more with less. This places great pressure on the decisionmaking process within the hospital, and has led, in Harris's words, to "a complicated system of rationing devices, uncodified rules, and subtle maneuvering, conducted in a sanctified atmosphere of 'life and death'."[5] These allocative rules are not based on price; the bargaining between the medical staff and the hospital administration does not take place within the traditional economic marketplace.

There seems to be a growing consensus that the costs of hospital care cannot be contained unless cost-minimizing forces are brought to bear on doctors as well as hospitals. There is no consensus as to *how* this should happen, and everyone shares a concern for the ethical implications of such cost-containment efforts.

In this essay, we will argue that attempts to reduce the doctor's use of medical resources through financial incentives do indeed raise ethical questions, but that the same questions are raised by systems that provide financial incentives to utilize resources. We will argue that discussions of cost containment have often oversimplified the ethical questions by not accounting for the complexity of the relationship among the benefits, risks, and costs of care, and by not recognizing the inadequacy of the evidence on what constitutes proper medical practice.

Benefits, Risks, and Costs

Only a few of the many complexities involved in evaluating the cost-effectiveness of medical practices[6] will be discussed here. Any decision by the doctor to use diagnostic or therapeutic resources may involve both potential benefits and risks of harm. When the expected potential benefits outweigh the expected potential harm, the expected net health effects are positive, and vice versa. In estimating the potential benefit or harm

[4] H Hiatt, *Protecting the medical commons: Who is responsible?* N Engl J Med (1975, 293:235-40).

[5] JE Harris, *The internal organization of hospitals: Some economic implications*, Bell J Econ (1977, 8:467-82).

[6] MC Weinstein, WB Stason, *Foundations of cost-effectiveness analysis for health and medical practices*, N Engl Med (1977, 296:716-21).

of any decision, the *probability* of the beneficial or harmful event must be multiplied by the (positive or negative) *value* of the event, when it occurs. Later we will argue that the *probability* of a benefit or harm can be measured objectively, and that it is important for such measurements to be made more often. However, the *value* of an event can be defined only subjectively and is inherently arbitrary.

In estimating the costs of a decision to use medical resources, one must consider both immediate expenditures and future costs that have been averted as a result of the medical action.

These observations highlight the importance of identifying the perspective from which benefits, risks, and costs are being considered. The patient, the doctor, the hospital, the insurance company, and society all have different perspectives and may place different values on the benefits and risks of harm. Also, the costs of a decision may be borne quite differently by different parties to a decision. Indeed, it may be that the costs involved in a medical decision are not all "medical" costs; for example, a screening program to detect a curable condition that produces learning disabilities in children will cost "medical" dollars to implement, but it will save future "educational" dollars; conversely, a school lunch program in an impoverished community will cost "educational" dollars to implement, but it may save future "medical" dollars by averting medical care for the consequences of malnutrition.

Although the perspective of the patient is inevitably unique, values may not be. Later we will cite examples of decisions that seem to be in everyone's interest—non-zero-sum games occurring within a "zero-sum society." Nevertheless, there clearly will be cases in which the patient's values run counter to those of society.

Many discussions of resource allocation and cost containment in medicine implicitly or explicitly assume that each medical resource the doctor might order will inevitably benefit the patient and will inevitably cost money—the question being whether the benefit is worth the cost.

If this assumption were true, every medical decision would contain the ethical questions implicit in considering such trade-offs. However, the assumption is invalid. It fails to recognize two important points: (1) More is not always better; (2) more can be less.

More Is Not Always Better. Medical care confers no guarantee of benefit. Indeed, the famous biochemist Lawrence Henderson once said: "Somewhere between 1910 and 1912 in this country . . . a random patient with a random disease consulting a doctor chosen at random had, for the first time in the history of mankind, a better than 50-50 chance of profiting from

the encounter."[7] The arguments of Illich notwithstanding, today the average patient's chances are considerably better than 50-50. But in the individual case, some medical actions may be wasteful: The benefits may not outweigh the harm (the net health effects may be zero or negative). With medical care, economists must be careful in applying the usual jargon; in some case, medical "goods and services" may be "bads and disservices."

There are numerous examples of waste in modern medicine, although we would not venture to estimate how much waste there is. Edwards reported that bacterial cultures from teaching hospital services were often ordered inappropriately and that culture results were not used in patient care in any "consistent logical" way.[8] Dixon reported that only 5% of clinical chemistry results at a teaching hospital were used in formulating diagnostic or therapeutic decisions.[9] Berwick found that of the large number of laboratory tests (40 per patient) ordered to evaluate children hospitalized with the "failure to thrive" syndrome, virtually none (0.8%) affected diagnosis or treatment; abnormal results were frequent, but not useful.[10] Even granting that an occasional child benefited greatly from such an evaluation, it should have seemed imperative to learn more about which children were most likely to benefit from such a diagnostic evaluation or which diagnostic tests were most likely to provide useful information. For the experience was certainly not benign. The children were removed from their homes, were placed in an unfamiliar and sometimes terrifying environment, and were subjected to tests that were often painful—almost always without a clear benefit. The responsible doctors were performing what they could only regard as a humanitarian function, but would the proverbial "man from Mars" make the same judgment?

Later, we will cite other data strongly suggesting that hospital care, surgery, and laboratory tests are sometimes wasted. Here, we will describe one further example: the use of medications, particularly antibiotics.

Three billion dollars a year and one-seventh of all hospital days are spent in the care of toxic reactions to medications, most of which are predictable and preventable.[11] Antibiotics are

---

[7] HL Blumgart, *Caring for the patient*, N Engl J Med (1965, 270:449-56).

[8] LD Edwards, S Levin, R Balagtas, *Ordering patterns and utilization of bacteriologic culture reports*, Arch Intern Med (1973, 132:678-82).

[9] RH Dixon, J Laszlo, *Utilization of clinical chemistry services by medical house staff*, Arch Intern Med (1974, 134:1064-67).

[10] DM Berwick, personal communication.

[11] See: KL Melmon, *Preventable drug reactions—causes and cures*, N Engl J Med (1971, 284:1361-68); JF Burnum, *Preventability of adverse drug reactions*, Ann Intern Med (1976, 85:80-81).

one of the most frequently prescribed types of medications. The overutilization of antibiotics occurs with hospitalized and ambulatory patients alike. Surveys of antibiotic use in community hospitals[12] and even in major university teaching hospitals[13] have indicated that 50-65% of antibiotics ordered either are not indicated at all or are being given incorrectly. For example, an estimated 3.3 million patients receive prophylactic antibiotics each year in American hospitals.[14] An extensive survey in Pennsylvania indicated that 80% of patients received prophylactic antibiotics for longer than could be justified medically.[15] Among ambulatory patients, it was estimated in one careful study that 25-40% of antibiotics were given unnecessarily.[16]

The use of antibiotics when no benefits can be expected does not simply waste money. It also introduces risks. Approximately 5% of hospitalized patients on antibiotics will suffer a toxic drug reaction.[17] In the ambulatory care setting, there are some disturbing examples of inappropriate and dangerous use. In their studies of the Tennessee Medicaid program, Ray and his colleagues found that 27% of the doctors prescribed tetracycline to children below the age of eight, a practice condemned by all authorities because it exposes children to the danger of permanently stained teeth.[18] They also found that 6% of the doctors prescribed chloramphenicol—a drug with potentially fatal side effects, and with virtually no clinical indications in the ambulatory patient.[19]

There is little evidence that doctors have any nefarious purpose (e.g., economic reward) in overutilizing antibiotics.

---

[12] *See:* WE Scheckler, JV Bennett, *Antibiotic usage in seven community hospitals*, JAMA (1970, 213:264-67); AW Roberts, JA Visconti, *The rational and irrational use of systemic antimicrobial drugs*, Am J Hosp Pharm (1972, 29:1054-60).

[13] CM Kunin, T Tupasi, WA Craig, *Use of antibiotics. A brief exposition of the problem and some tentative solutions*, Ann Intern Med (1973, 79:555-60); M Castle, CM Wilfert, TR Cate, S Osterhout, *Antibiotic use at Duke University Medical Center*, JAMA (1977, 237:2819-22).

[14] M Shapiro, TR Townsend, B Rosner, EH Kass, *Use of antimicrobial drugs in general hospitals. Patterns of prophylaxis*, N Engl J Med (1979, 301:351-55).

[15] *Ibid.*

[16] RH Brook, KN Williams, *Effect of medical care review on the use of injections. A study of the New Mexico experimental medical care review organiztion*, Ann Intern Med (1976, 85:509-15).

[17] LE Cluff, *The prescribing habits of physicians*, Hosp Pract (1967, 2:100-04).

[18] WA Ray, CF Federspiel, W Schaffner, *Prescribing of tetracycline to children less than 8 years old. A two-year epidemiologic study among ambulatory Tennessee Medicaid Recipients*, JAMA (1977, 237:2069-2101).

[19] WA Ray, CF Federspiel, W Schaffner, *Prescribing of chloramphenicol in ambulatory practice. An epidemiologic study among Tennessee Medicaid recipients*, Ann Intern Med (1976, 84:266-70).

The practice is probably well-intentioned, if uninformed. Patients often do expect or demand antibiotics. However, few examples in medicine better demonstrate these points: Beneficial forms of therapy are not beneficial in every patient, only in some patients; beneficial forms of therapy often contain risks, as well; everything the doctor tries to do *for* the patient he or she may also be doing *to* the patient; more is not always better.

More Can Be Less. Just as an investment in medical care does not necessarily buy benefits, it does not necessarily involve an additional cost. In fact, an additional dollar investment today may save future dollars, resulting in a net saving of money. For example, there is little doubt that the initial cost of polio immunization, or of hypertension treatment, or of lead-screening programs in high-risk communities[20] can produce substantial net dollar savings through averting future medical and social services. Positive health effects and dollar savings can be achieved simultaneously. More can be less.

Frictionless Resource Allocation. In any society, the allocation of resources is a difficult and fractious process. It could be otherwise if three conditionss were met:

(1) If society had accurate information about the outcomes and associated probabilities of all medical actions;

(2) if all members of society valued the various outcomes in the same way, and their values were known;

(3) if society had reached a consensus on how to compare one person's well-being with another's, and on how to determine a fair distribution of benefits and costs.

If these three conditions could be met, then society would not have to establish systems of economic incentives and disincentives, nor would it have to worry about social and psychological pressures distorting medical decisions. When the risks of harm outweigh the benefits—when an action is wasteful and produces net dollar costs as well, it would not be taken; when the benefits outweigh the risks and costs are saved as well, the action would be taken; and when the action is associated with both net positive benefits and dollar costs, an allocation decision would be made on the basis of the consensus of values.

It is obvious that a consensus of values does not now exist in our society, and it is doubtful that one ever will. Indeed, it seems more likely that our society will reduce its resource allocation problems by achieving technological breakthroughs that exponentially amplify available resources (e.g., energy).

The point we would stress, however, is that the allocation of medical resources in our society is made difficult not only

[20] DM Berwick, AL Komaroff, *Cost-effectiveness of lead screening*, N Engl J Med (1982, 306:1392-98).

because of the absence of a consensus of values. It is difficult also because society does not have the information upon which rational decisions can be based, information about the probabilities of benefits and risks and about costs, information that identifies practices in everyone's interest and those in no one's interest, information necessary (though not sufficient) for defining "proper medical practice." At least in principle, the problem of inadequate information is remediable.

Proper Medical Practice and a Doctor's Decisions

"Proper Medical Practice" Is Ill-Defined. Voltaire once remarked that there were "no sects in geometry." Presumably, he was referring to the fact that the science of geometry employs an explicit set of assumptions and a terminology free of redundancy and ambiguity, makes measurements of a nature allowing great precision, and is based in theorems that can all be proven. For lack of these same virtues, there are sects in medicine.

The elements of a medical decision can be divided theoretically into objective and subjective components. The subjective components—e.g., different values given to the same outcome, a fear of censure and a desire for praise, a fear of economic penalties, and a desire for rewards—are ubiquitous.

Discussions of cost containment in medicine often seem to assume that the objective components associated with a particular medical action—the probabilities of different benefits and harms, and the dollar costs—are stable, established, and well-known. From this assumption it would follow that variability in medical practice could be attributed entirely to subjective factors.

However, the assumption is wrong. More often than not, medical science has not provided the doctor with information regarding the probabilities of different benefits and harmful events, and medical managers and planners have not calculated actual costs (immediate or induced). As a consequence, our ignorance about the values of the objective components of medical decisions, coupled with the variable nature of the subjective components of decisions, obscures the definition of "proper medical practice."

This argument may seem surprising, given the remarkable strides made in the biomedical sciences during this century. It is true that we are beginning to have a fundamental understanding of normal biology and the pathogenesis of disease. But that understanding, no matter how profound, cannot guarantee answers to questions about proper medical practice.

Consider, for example, three scientific achievements of enormous importance: the discoveries of insulin, of penicillin, and of the molecular structure of normal and variant hemo-

globins. Almost 60 years later, physicians still cannot identify *which* diabetic patients need insulin; great controversy persists about the importance of blood sugar control in that large group of diabetics who remain asymptomatic and free of ketoacidosis. Almost 40 years after the availability of penicillin, we do not know *which* patients with a sore throat should receive this antibiotic; the risks and costs of penicillin must be considered in light of the inaccuracy of the throat culture and the diminishing virulence of streptococcal infection.[21]

Almost 40 years after discovering the molecular basis of sickle hemoglobin, we cannot answer many questions about proper practice. When the patient with sickle cell disease arrives at the emergency room with a fever, for example, when should he be tested for an occult infection? Always, even though sickle cell crisis alone often produces fever? If so, which tests for infection should be used: urine culture, blood culture, bone marrow culture, spinal puncture, biliary ultrasound, bone scan, gallium scan, computed tomography? Ask four physicians and you will get (at least) four opinions. Ask the scientists who pioneered in the discoveries of hemoglobin structure, confer with Perutz or make a pilgrimage to Pauling, and you will get no opinion. An understanding of the pathogenesis of disease, or the discovery of a cure for a disease, still leaves unanswered many questions about the proper practice.

For to answer such questions, we need a new science—a science of clinical practice—in which information about the probabilities and the magnitudes of the relevant benefits, harm, and costs are determined empirically, and in which the information is synthesized for and readily accessible to the doctor and the patient.

Necessary, Discretionary, and Wasteful Care. Rather than thinking of medical care as representing either proper or improper medical practice, we find it more useful to describe care as "necessary," "discretionary," or "wasteful."

Necessary care includes care that virtually all doctors would agree is indicated because the benefits greatly outweigh the risks, even considering the range of values different individuals would place on the outcome. Most true medical emergencies are a good example. There are few economic constraints on the giving of emergency care in the United State, at least once the patient has reached the hospital. Even if the efficacy of the practice has not been demonstrated by a formal scientific study, there is a near-unanimous consensus among doctors about its necessity. Only with (temporarily) life-sustaining care to the terminally ill is rationing of necessary care

[21] AL Komaroff, *A management strategy for sore throat*, JAMA (1978, 239:1429-32).

ever considered. Necessary care probably constitutes only a small fraction of all medical care given in the United States.

Discretionary care is care in which the relative weight of benefit and harm is uncertain—either because different subjective values may tip the balance in either direction, or because objective data regarding the probabilities of harm and benefit is not well established. With discretionary care, it is likely that different doctors and patients will make quite different decisions under the same circumstances.

Most medical care is discretionary. And it is with discretionary care that traditional economic incentives are most likely to affect the behavior of both doctors and patients, despite the forces that interfere with the free play of the medical market. When the doctor is allowed a wide range of "acceptable" behavior, it is easier for subjective (economic, psychological, and social) factors to influence decisions about "supply." And when the patient perceives care as discretionary, the price elasticity of demand is also likely to be at its greatest.

Wasteful care, as described earlier, occurs when the risks of an action can be expected to equal or exceed the benefits, when the net health effects are zero or negative. To be called wasteful, scientific evidence or a general consensus must have established that this is the case. We do not know how much waste there is in medical care.

Pressures On The Doctor's Resource Allocation Decisions

Pressures for use. The economic pressures on the doctor to utilize resources in the fee-for-service system are widely recognized. The market does not work well in medical care. The patient does not have the information or the freedom of choice of a true consumer.[22] Indeed, the assumption that health care is a basic right, coupled with the urgency felt by the ill patient tends to create an entitlement—tends to transform the patient from a "consumer" into a "beneficiary."[23] The role of the doctor in defining a patient as "sick" and in need of certain services transforms the doctor into a hybrid "provider-consumer."

Not all of the forces that encourage a doctor to utilize medical resources are economic. Psychological pressures to know as much as possible about and to do as much as possible for the patient—Fuchs's "technological imperative"[24]—have a substantial impact on the use of diagnostic and therapeutic

---

[22] KJ Arrow, *Problems of resource allocation in United States medical care,* EXPERIENTIA (1972, 17[suppl]:392-410).
[23] K Brewster, *Health at any price?* J ROY SOC MED (1979, 72:719-22).
[24] VR Fuchs, WHO SHALL LIVE? HEALTH, ECONOMICS, AND SOCIAL CHOICE (New York: Basic Books, Inc., 1974).

resources. A desire for prestige, a wish to avoid malpractice charges, and perhaps most important, a sense of the action expected by his or her professional peers and by the patient—all influence the doctor's decisions. "Inaction" is equated with indecision or ignorance. The feeling is expressed most parsimoniously by that axiom known to every surgical intern: "When in doubt, cut it out."

Pressures for rationing. Many people have expressed the concern that increased efforts to conserve hospital care will lead to rationing of medical resources. The assumption seems to be that medical resources are not now rationed nor have they been in the past. Indeed, our just-concluded recitation of the pressures to utilize resources would seem to support such an assumption. However, the assumption is wrong. Rationing has always occurred in medical care.

Before defending this proposition, we first must define "rationing." We are not referring to rationing in the strictest sense: a formal and explicit system to allocate equitably a resource that has been identified as "scarce." There are only a few such examples in medicine (e.g., kidney dialysis in its early years). However, there certainly are mechanisms formally and explicitly designed to reduce demand for all resources, scarce or otherwise.

Since its inception the United States has rationed health care. Before health insurance was widespread, rationing was by price. Since health insurance has become commonplace, rationing has persisted in the form of deductibles, coinsurance, or benefit packages that exclude psychiatric treatment, dental care, eyeglasses, medications, or cosmetic surgery. Rationing is also achieved by access barriers that effectively reduce demand: queues to see a doctor for nonemergency conditions, costs in time and dollars to travel to the doctor, and queues after arriving to see the doctor.

Even in the most high-powered hospitals there are always queues. Certain laboratory tests and certain treatments are simply not available on the weekends or in the evenings, except when the doctor can convince someone that a critical emergency exists. Care may be delayed for reasons of convenience, especially for procedures that require professional skill—even when the patient's condition is quite serious. For example, it is probably true that endoscopy, radiologic contrast studies, and angiography are not performed as promptly on patients with gastrointestinal bleeding when such patients are admitted to the hospital in the evening or on the weekend rather than during the working day.

Indeed, in virtually every encounter with the patient the doctor rations time. We rarely are as thorough in taking a medical history, performing a physical examination, or considering the diagnostic and therapeutic options as we know we could be.

In fact, this sense of having been incomplete may be a stimulus to increase diagnostic tests. Griner demonstrated that doctors were much less likely to obtain a diagnostic test if they had to perform the test themselves;[25] like most human beings, doctors are more generous with other people's time than with their own.

Rationing also occurs in another form, one usually overlooked in discussions of the pressures on doctors to utilize medical resources: Sometimes doctors try to save their patients' money. This occurs when the doctor judges that a discretionary test or treatment is of marginal value, and knows the patient will have to pay some or all of the cost out-of-pocket. In such cases, the doctor may feel a responsibility to protect the patient from economic as well as physical adversity. There are no data by which this assertion can be defended, but we believe it to be true and reasonably common. It is more likely to occur in the over one billion visits for ambulatory care each year than in the hospital, because of differential insurance coverage. But since most people "covered" by hospital insurance still incur some risk-sharing through coinsurance, the concern is not moot even in the hospitalized patient.

Variability and Misutilization in Medical Practice

Elsewhere we have argued that doctors collect and interpret medical data with surprising variability and inaccuracy.[26] In this essay, we argue that variability is characteristic also of medical practice, of decisions about the use of medical resources. Before considering the evidence, a word of caution is necessary. Virtually all of the studies cited below are limited by lack of enough clinical information to confidently conclude that the differences observed represent differences in doctors' judgments rather than differences in the nature and severity of patients' illnesses. Furthermore, none of the studies is able to prove that any particular factors explain the observed differences. Nevertheless, the weight of the evidence suggests that much of the variability cannot be explained only by differences in illnesses. As we have argued earlier, it is very likely that subjective components of a medical decision—social, psychological, and economic factors—explain a great deal of the observed variability, as does a lack of information about the objective components of the decision.

---

[25] PF Griner, B Liptzin, *Use of the laboratory in a teaching hospital. Implications for patient care, education, and hospital costs*, Ann Intern Med (1971, 75:157-63).

[26] AL Komaroff, *The variability and inaccuracy of medical data*, Proc of the IEEE (1979, 67:1196-1207).

Hospitalization. The doctor decides that the patient requires hospitalization and determines the length of stay. Since charges for hospital days constitute the largest part of the hospital bill, these decisions have considerable economic importance.

A number of studies have concluded that between 5-20% of acute care hospital days are wasteful, offering no benefit.[27] These conclusions have been based on personal judgments, and only those that have used explicit criteria to supplement implicit judgments have demonstrated adequate inter-judge reliability.[28] As Lust has summarized, many studies also have demonstrated that the utilization of hospital days is markedly lower with various forms of prepaid insurance than with the traditional fee-for-service system.[29]

In this essay, we will concentrate on studies that describe variability of practice without making judgments about appropriateness of hospital care, studies that demonstrate variability within the same insurance scheme—i.e., variability not attributable to differences between prepaid and fee-for-service insurance schemes. A number of studies have demonstrated such variability.[30] Only a few will be reviewed in any detail.

---

[27] *See:* PM Gertman, JD Restuccia, *The appropriateness evaluation protocol: A technique for assessing unnecessary days of hospital care,* Med Care (1981, 19:855-71); JG Zimmer, *Length of stay and hospital bed misutilization,* Med Care (1974, 12:453-62); JD Restuccia, DC Holloway, *Barriers to appropriate utilization of an acute facility,* Med Care (1976, 14:559-73); RL Berg, FE Browning, SL Crump, W Wenkert, *Bed utilization studies for community planning,* JAMA (1969, 207:2411-13); FH Lovejoy, Jr, JM Carper, CA Janeway, J Kosa, *Unnecessary and preventable hospitalizations: Report on an internal audit,* Med Care (1971, 79:868-72); GE Vaillant, LN Shapiro, PP Schmitt, *Psychological motives for medical hospitalization,* JAMA (1970, 214:1661-65); JM MacKintosh, T McKeown, FN Garratt, *An examination of the need for hospital admission,* Lancet (1961, 1:815-18); AI Mushlin, FA Appel, *Extramedical factors in the decision to hospitalize medical patients,* Am J Public Health, (1976, 66:170-72).

[28] *See:* Gertman and Restuccia, Restuccia and Holloway, *supra* note 27.

[29] HS Luft, *How do health-maintenance organizations achieve their "savings"?* N Engl J Med (1978, 298:1336-43).

[30] *See:* SA Schroeder, DS O'Leary, *Differences in laboratory use and length of stay between university and community hospitals,* J Med Educ (1977, 52:418-20); S Lichtner, M Pflanz, *Appendectomy in the Federal Republic of Germany: Epidemiology and medical care patterns,* Med Care (1971, 9:311-30); E Jonsson, D Neuhauser, *Hospital staffing ratios in the United States and Sweden,* Inquiry (1975, 12:128-37); H Sherman, M Flatley, *Dissecting the hospital stay. A method for studying patient staging in hospitals,* Med Care (1980, 18:715-30); DR Sheeley, H Sherman, *Conservation in hospital resource use: Treatment of pneumonias. An investigation of care in four hospitals over the past decade,* Ann Intern Med (1976, 85:648-52); MA

Individual physicians working within the same medical setting demonstrate considerable variability in hospitalization practices. Lyle and colleagues[31] studied a group practice of seven internists. The physicians' patients were similar in age, sex, and diagnosis. Yet there were differences of up to twofold in the frequency of hospital admission, adjusted by diagnostic category, as well as for the average length of stay (LOS). Each physician tended to have a consistent LOS (high, average, low) across all diagnostic categories. The evidence suggests that, for patients with comparable illnesses, the doctors differed both in their inclination to hospitalize patients and in their decisions about LOS.

Much of the evidence for variable medical practice, including the use of hospital beds, comes from Britain. This is interesting, given the often expressed assumption that socialization leads to homogenization of social services.

West and Carey[32] compared rates of admission (per 1000 population per year) for appendicitis for 17 health districts in Wales. The highest age-adjusted rate was 2.04, and the lowest was 0.64. Admissions were more likely to occur during the week than on weekends. The authors concluded that the incidence of appendicitis could not vary that greatly in different parts of Wales, and they attributed the differences in admission rates to hospital policies and availability of beds and physicians.

In a study of several hospitals throughout England and Wales, Heasman and Carstairs found a considerable variation in LOS from one hospital to the next, for the same medical and surgical conditions.[33] For acute myocardial infarction the range of median LOS was 10-36 days; for peptic ulcer, 6-26 days; for hernia repair, 2-12 days; for lens extraction, 4-18 days; for excision of the semilunar cartilage, 3-21 days; and for hysterectomy, 3-18 days. The variability characterized both the preoperative and postoperative period. Similar variability was demonstrated in another study by Heasman.[34]

Surgery. Studies in the United States, Canada, Britain, Europe, and Japan have demonstrated considerable variability in

---

Heasman, V Carstairs, *Inpatient management: Variations in some aspects of practice in Scotland*, Br Med J (1971, 1:495-98); PF Griner, *Treatment of acute pulmonary edema: Conventional or intensive care?* Ann Intern Med (1972, 77:501-06).

[31] CB Lyle, DS Citron, WV Sugg, *Cost of medical care in a practice of internal medicine. A study in a group of seven internists*, Ann Intern Med (1974, 81:1-6).

[32] RR West, MJ Carey, *Variation in rates of hospital admission for appendicitis in Wales*, Br Med J (1978, 1:1662-64).

[33] *See:* Heasman and Carstairs, *supra* note 30.

[34] MA Heasman, *How long in hospital? A study in variation in duration of stay for two common surgical conditions*, Lancet (1964, 2:539-41).

surgery rates. Even when the insurance and payment schemes are the same, doctors decide differently about the need for surgery.

Perhaps the most dramatic evidence of variability in surgical rates in the United States has come from the work of Wennberg and Gittlesohn. In 1969, these authors examined the variation in resources and in utilization of medical services in 13 hospital catchment areas of Vermont.[35] Total age-adjusted surgery rates per 10,000 of specific frequently performed surgical procedures varied strikingly among the 13 geographical areas, as shown in Table 1. Similar patterns were noted in a later study in Maine.[36] The authors demonstrated a significant correlation between the total surgery rate and the supply of surgeons in the different geographical areas. The authors concluded that "variation among geographical areas is the rule," and that the supply of beds and surgeons had a greater effect than consumer behavior on surgical rates, since neighboring areas with apparently homogeneous populations but different numbers of beds and surgeons had vastly different rates.

TABLE 1: Variability in Surgical Rates, Vermont, 1969*

Procedure	Highest Rate	Lowest Rate	Ratio: Highest/ Lowest
Tonsillectomy	151	13	11.6
Appendectomy	32	10	3.2
Cholecystectomy	57	17	3.4
Hysterectomy	60	10	6.0
Uterine dilatation and curettage	141	30	4.7
Varicose veins	28	6	4.7
TOTAL	689	360	1.9

*Rates in patients of all ages, per 10,000 population.

SOURCE: Adapted from Wennberg J, Gittlesohn A, *Small area variations in health care delivery*, SCIENCE (1973, 182:1105 [Table 3]).

[35] J Wennberg, A Gittelsohn, *Small area variations in health care delivery*, SCIENCE (1973, 182:1102-08).

[36] JE Wennberg, A Gittelsohn, *Health care delivery in Maine. I: Patterns of use of common surgical procedures*, J MAINE MED ASSOC, (1975, 66:123-49).

Lewis found similar variations in surgical rates in Kansas.[37] Rates of surgery per 10,000 enrollees in a medical insurance plan varied from 153-432 for tonsillectomy, from 15-61 for appendectomy, and from 12-42 for cholecystectomy. Rates of surgery in a particular area were significantly related to the number of physicians who performed surgery and to the availability of hospital beds and surgeons. Lewis concluded that differences in the frequency or severity of different diseases probably were not responsible for the geographic variation in surgical rates.

Are these conclusions justified? As discussed earlier, the data available were not adequate to conclude with certainty that differences in the nature and severity of the patients' illnesses did not account for the differences in the doctors' decisions about surgery. However, this question was addressed indirectly in a subsequent study by Wennberg and his colleagues.[38]

In that study, Wennberg measured the changes in the rate of tonsillectomy after the findings of the first study were reported to Vermont surgeons. As shown in Table 2, the results were dramatic. The area with the highest rate in 1969 changed its practices decisively, dropping its rate by 89%. While the ratio of the highest to the lowest rate of tonsillectomy in the different geographic areas had been 13.5:1 in 1969, by 1973 the ratio was 4.5:1. The overall rate of tonsillectomy in Vermont dropped 46% from 1969 to 1973, a much greater change than occurred nationwide. It seems much more plausible to attribute a change of that magnitude over such a short time to a change in medical practice rather than to a change in patient population. And this suggests that the great inter-area variability noted in 1969 was also best explained by variability in medical practice.

Although most studies,[39] like those of Wennberg and Lewis, have suggested that surgical rates are related to the supply of surgeons, at least two studies have failed to find a convincing relationship.[40] Although some studies have concluded that sur-

---

[37] CE Lewis, *Variations in the incidence of surgery*, N ENGL J MED (1969, 281:880-84).

[38] JE Wennberg, L Blowers, R Parker, AM Gittelsohn, *Changes in tonsillectomy rates associated with feedback and review*, PEDIATRICS (1977, 59:821-26).

[39] *See:* Wennberg and Gittelsohn, *supra* note 35, 36; Lewis, *supra* note 37. *Also see:* JP Bunker, *A comparison of operations and surgeons in the United States and in England and Wales*, N ENGL J MED (1970, 282:135-44); E Vayda, *A comparison of surgical rates in Canada and in England and Wales*, N ENGL J MED (1973, 289:1224-29); E Vayda, GD Anderson, *Comparison of provincial surgical rates in 1968*, CAN J SURG (1975, 18:18-26); Y Yoshida, K Yoshida, *The high rate of appendectomy in Japan*, MED CARE (1976, 14:950-57).

[40] *See:* NP Roos, LL Roos, Jr., PD Henteleff, *Elective surgical rates—Do high rates mean lower standards?* N ENGL J MED (1977, 297:360-65); JP LoGerfo, RA Efird, PK Diehr, WC Richardson, *Rates of surgical care in*

TABLE 2: Change in Tonsillectomy Rates, Vermont 1969-73*

Geographic Area	1969 Rate	1973 Rate	Change: 1969-73 (%)
Highest 1969 rate	444	50	−89
Lowest 1969 rate	33	27	−18
Average 1969 rate	129	70	−46
U.S. average rates	136	117	−14

*Rates per 10,000 children, 14 years of age or younger.

SOURCE: Adapted from Wennberg *et al.*, *Changes in tonsillectomy rates associated with feedback and review*, PEDIATRICS (1977, 59:824 [Table 3]).

gical rates are related to the availability of hospital beds,[41] other studies have failed to show such a relationship.[42] Patient expectations may also play a role in the decision to perform surgery. Roos and colleagues[43] reported that, in Britain, tonsillectomy was more frequent among children who had also been circumcised and whose mothers had suffered from gynecologic conditions. Lichtner and Pflanz[44] suggested that the reason appendectomy was 3-4 times more common in white-collar workers than in blue-collar workers was that surgeons were less likely to risk the consequences of failing to perform surgery in patients with appendicitis when the patients were of higher social standing.

In summary, the apparent variability in surgical practices may be accounted for entirely by clinical differences among patient populations with the same diagnoses. The evidence is not conclusive that surgical rates vary because of variability in medical practice, or that variable practices can be explained largely either by subjective factors (economic, social, and psychological), or by ignorance of objective factors (e.g., the probabilities of benefit or harm). However, one can be forgiven for concluding that this is the case.

---

prepaid group practices and the independent setting. What are the reasons for the differences? MED CARE (1979, 17:1-9).

[41] *See:* Lewis, *supra* note 37; Bunker, *supra* note 39.

[42] DB Dutton, *Patterns of ambulatory health care in five different delivery systems*, MED CARE (1979, 17:221-43).

[43] *See:* Roos *et al.*, *supra* note 40.

[44] *See:* Lichtner and Pflanz, *supra* note 30.

Diagnostic and Therapeutic Technologies. The number of diagnostic tests and procedures is growing at a rate of about 15% per year; typically, tests account for about 25% of the hospital bill.[45] The growth in laboratory tests seems to be a general phenomenon, found with many different illnesses. Scitovsky demonstrated a two- to sixfold rise in the use of laboratory tests for selected illnesses at one medical center over the 20-year period between 1951 and 1971.[46]

There is considerable evidence from practices in the United States and elsewhere of marked variability in the utilization of laboratory tests.[47] We shall discuss only a few.

In perhaps the most dramatic example, Schroeder and colleagues[48] performed an audit of the cost of laboratory tests ordered during a three-month period by 33 salaried faculty internists at a teaching hospital ambulatory clinic. The average aggregate annual test costs per patient ranged from $10-$169 for different physicians—a 17-fold difference. Given that variability of this degree has not been reported from other studies of test ordering, and that the severity of disease was not measured carefully in this study, it is likely that some of the variability can be explained by differences in case mix.

Salaried internists of the Kaiser-Portland Health Plan ordered from less than 1 to over 2.5 tests per office visit, a nearly threefold difference.[49] The practices of the internists were similar in terms of diagnostic mix and case load, and the ranking of

---

[45] *See:* Gertman and Restuccia, *supra* note 27. *Also see:* HV Fineberg, *Clinical chemistries: The high cost of low cost diagnostic tests*, in: S Altman, R Blendon (eds.), MEDICAL TECHNOLOGY: THE CULPRIT BEHIND HEALTH CARE COSTS? DHEW Publication no. (PHS) 79-3216 (Washington, DC: Government Printing Office, 1979:144-65); DM Smith, SD Roberts, TL Cross, *Components of the costs of care for medicine patients at an urban care center*, CLIN RES (1979, 27:284A).

[46] *See:* Scitovsky, *supra* note 2.

[47] *See:* M Daniels, SA Schroeder, *Variation among physicians in use of laboratory tests. II. Relation to clinical productivity and outcomes of care*, MED CARE (1977, 15:482-87); SA Schroeder, A Schliftman, TE Piemme, *Variation among physicians in use of laboratory tests: Relation to quality of care*, MED CARE (1974, 12:709-13); DF Hardwick, P Vertinsky, RT Barth, VF Mitchell, M Bernstein, I Vertinsky, *Clinical styles and motivation: A study of laboratory test use*, MED CARE (1975, 13:397-408); JR Hampton, MJG Harrison, JRA Mitchell, JS Prichard, C Seymour, *Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients*, BR MED J (1975, 2:486-89).

[48] SA Schroeder, K Kenders, JK Cooper, TE Piemme, *Use of laboratory tests and pharmaceuticals. Variation among physicians and effect of cost audit on subsequent use*, JAMA (1973, 225:969-73).

[49] DK Freeborn, D Baer, MR Greenlick, JW Bailey, *Determinants of medical care utilization: Physicians' use of laboratory services*, AM J PUBLIC HEALTH (1972, 62:846-53).

the physicians in terms of laboratory use remained consistent over several years. The authors tried to relate the use of laboratory tests to certain physician characteristics: age, board certification, recent graduation from medical school, geographic location of medical school, length of time on the Kaiser staff, leadership role in Kaiser organization, and utilization of laboratory tests by an important "role model"—the chief of service. Of all these, only the last was significantly associated with the test utilization patterns of individual physicians.

The above examples are from ambulatory care practice. Variability in the use of nonsurgical technologies among hospitalized patients in the United States has not been examined frequently. In a recent study at the Johns Hopkins Hospital, investigators found a greater than twofold variation in the aggregate charges of ancillary resources (laboratory and radiologic studies and pharmacy services) used per day, when different medical units with apparently the same patient mix were compared. There were greater than threefold differences in laboratory utilization among interns, when dealing with approximately the same patient mix.[50]

Ashley and colleagues[51] examined the use of pathology and radiology services in England and Wales, according to 15 common diagnostic categories. Weighted units of use were calcuated according to technician time required. Among eight hospitals studied, very wide variations were found in the average number of pathology service units per patient: For hernia repair, the range was 2-53 units; for prostatectomy, 78-215 units; for coronary thrombosis, 48-116 units. Wide variation was also seen in the average number of radiology test units: For chronic bronchitis, the range was 0.9-2.2 units; for peptic ulcer, 4.0-9.5 units; for hernia repair, 0.1-2.4 units; for prostatectomy, 2.4-8.2 units. The authors examined length of stay, age, and presence of complications, but could not explain laboratory utilization differences on the basis of differences in case mix.

In each of the above examples of diagnostic test ordering, it should be noted that no differential economic pressures existed. Rather, the differences seem more likely to be accounted for by other subjective factors, social and psychological.

As for therapeutic technology (other than surgery), one example of variability is particularly striking: the use of renal dialysis. In 14 Western European countries (and Israel), dialysis rates range from 6-144 per million population. In the United States in 1979, the national average rate was 205, ranging from

---

[50] RJ Johns, BI Blum, *The use of clinical information systems to control cost as well as to improve care*, Trans Am Clin Climatol Assoc (1979, 90:140-52).

[51] JSA Ashley, P Pasker, JC Beresford, *How much clinical investigation?* Lancet (1972, 1:890-93).

983 (District of Columbia) to 20 (Wyoming). In the most populous and representative states, the rates ranged from 119-282. Two eminent physicians who reviewed this data could not conclude that variable economic incentives explained the differences. They did conclude, however, that the differences were not likely due to differences in the prevalence of end-stage renal disease, stating that "in some states we are probably not treating all those who might benefit; in others we may be treating patients who cannot benefit."[52]

Interventions

Some experience has accumulated with four different strategies to influence the practice of doctors, as summarized recently by Myers and Schroeder[53]: education, audit with feedback, restrictions or rationing, and positive incentives.

Education. With regard to laboratory testing, a number of studies have demonstrated that—at least in the short-term—educational programs can result in the reduction of wasteful test ordering.[54] The educational sessions have concentrated on presenting evidence about the value of a test, as well as providing knowledge about the cost of the test. Indeed, there is ample evidence that medical students, interns, residents, and even medical faculty are equally uninformed about the prices of the tests and treatments they order.[55]

One study of this question is worth special mention. At the Strong Memorial Hospital in Rochester, New York, Griner and his colleagues conducted an extensive educational effort over a

[52] AS Relman, D Rennie, *Treatment of end-stage renal disease. Free but not equal*, N Engl J Med (1980, 303:996-98).

[53] LP Myers, SA Schroeder, *Physician use of services for the hospitalized patient: A review with implications for cost containment*, Milbank Mem Fund Q (1981, 4:481-507).

[54] See: JM Eisenberg, *An educational program to modify laboratory use by house staff*, J Med Educ (1977, 52:578-81); RL Rhyne, SH Gelbach, *Effects of an educational feedback strategy on physician utilization of thyroid function panels*, J Fam Pract (1979, 8:1003-07); D Henderson, R D'Alessandri, B Westfall, *et al.*, *Hospital cost containment: A little knowledge helps*, Clin Res (1979, 27:279A); KL McCoy, *The Providence Hospital blood conservation program*, Transfusion (1962, 2:3-6); PD Mintz, K Lauenstein, J Hume, JB Henry, *Expected hemotherapy in elective surgery: A follow-up*, JAMA (1978, 239:623-25).

[55] See: SP Kelly, *Physicians' knowledge of hospital costs*, J Fam Practice (1978, 6:171-72); SJ Dresnick, WI Roth, BS Linn, TC Pratt, A Blum, *The physician's role in the cost-containment problem*, JAMA (1979, 241:1606-09); JK Skipper, G Smith, JL Mulligan, ML Garg, *Medical students' unfamiliarity with the cost of diagnostic tests*, J Med Educ (1975, 50:683-84); LR Kirkland, *The physician and cost containment*, JAMA (1979, 242:1032).

seven-year period to reduce the use of wasteful hematology, chemistry, microbiology, and chest x-ray examinations.[56] While test use increased during the seven years, the rate of increase was considerably less than the national average, and less than that seen at a nearby affiliated teaching hospital. The great strength of this study is its examination of behavior over a far longer period of time than any similar study investigating the impact of an educational program. The weakness of the study is that it lacked a concurrent comparison group for the study period.

Medical practice is also "educated" by the results of formal studies of practice. With several medical and surgical conditions, when formal experiments have demonstrated that earlier discharge from the hospital could be achieved without sacrificing the quality of care, the medical community has changed its practice. A good example is myocardial infarction. In the late 1960s, the typical hospitalization for this condition lasted three weeks. Since then, a series of randomized clinical studies have shown that many patients require no longer than 7-14 days,[57] and a sizable fraction may not require hospitalization at all.[58] Probably because of this evidence, the average hospitalization for patients in 1981 was around two weeks in most communities.

Audit with Feedback. The approach of audit with feedback has generally been more successful than education alone. The approach has been tried with "big ticket" technologies such as surgery, and with "little ticket" technologies such as diagnostic laboratory tests. Over the past 30 years, hospital tissue review committees appear to have had a dramatic influence on the rate of surgery.[59] As described earlier, audit and feedback probably

[56] PF Griner, Medical House Staff, Strong Memorial Hospital. *Use of laboratory tests in a teaching hospital: Long term trends: Reductions in use and relative cost,* Ann Intern Med (1979, 90:243-48).

[57] See: BM Groden, A Allison, GB Shaw, *Management of myocardial infarction: The effect of early mobilisation,* Scott Med J (1967, 12:435-36); JE Harpur, RJ Kellett, WT Connor, *et al., Controlled trial of early mobilization and discharge from hospital in uncomplicated myocardial infarction,* Lancet (1971, 2:1331-34); AM Hutter, Jr., VW Sidel, KI Shine, *et al., Early hospital discharge after myocardial infarction,* N Engl J Med (1973, 288:1141-44); MJ Hayes, GK Morris, JR Hampton, *Comparison of mobilization after two and nine days in uncomplicateD myocardial infarction,* Br Med J (1974, 3:10-13).

[58] See: HG Mather, DC Morgan, NG Pearson, KL Read, DB Shaw, GR Steed, MG Thorne, CJ Lawrence, IS Riley, *Myocardial infarction: A comparison between home and hospital care for patients,* Br Med J (1976, 1:925-29); HG Mather, NG Pearson, KLQ Read, DB Shaw, GR Steed, MG Thorne, S Jones, CJ Guerrier, CD Eraut, PM McHugh, NR Chowdhury, MH Jafary, TJ Wallace, *Acute myocardial infarction: Home and hospital treatment,* Br Med J (1971, 3:334-38).

[59] PA Lembcke, *Medical auditing by scientific methods,* JAMA (1956, 162:646-55).

accounted for a dramatic decline in tonsillectomy rates in Vermont.[60]

Another impressive example is Dyck's attempt to reduce wasteful hysterectomies in Saskatchewan.[61] After ongoing audit was instituted to assess the justification for each hysterectomy according to explicit criteria, the rates in the three largest cities dropped from 1258-522 per 100,000 women, from 755-628, and from 499-356, respectively. The percentage of "unjustified" hysterectomies (according to the audit criteria) dropped markedly in the three cities, from 50% to 15%, from 17% to 7%, and from 20% to 4%, respectively.

Several studies have demonstrated that audit with feedback can reduce the wasteful utilization of antibiotics.[62] All of these audits have been brief; it is probable that constant or recurrent audit is required.

As for laboratory tests, Schroeder and his colleagues were able to reduce aggregate annual laboratory charges by 29% in a group of 33 faculty internists by providing information about both the charges generated by their test-ordering practices and their own performance compared with that of their colleagues.[63] Others have reported a similar positive experience, at least in the short term.[64]

The most rigorous study of the impact of audit with feedback was conducted by Martin and his colleagues.[65] They conducted a prospective study in which medical residents were randomly allocated to experimental and control groups. The group whose medical records were audited regularly by senior physicians ordered almost 50% fewer laboratory tests than a control group, a statistically and clinically significant difference. The great strength of this study was its ability to control

---

[60] *See:* Wennberg *et al.*, *supra* note 38.

[61] FJ Dyck, FA Murphy, JK Murphy, DA Road, MS Boyd, E Osborne, D DeVlieger, B Korchinski, C Ripley, AT Bromley, PB Innes, *Effect of surveillance on the number of hysterectomies in the province of Saskatchewan*, N Engl J Med (1977, 296:1326-28).

[62] *See:* Brook and Williams, *supra* note 16. *Also see:* WA Craig, SJ Uman, WR Shaw, V Ramgopal, LL Eagan, ET Leopold, *Hospital use of antimicrobial drugs. Survey at 19 hospitals and results of antimicrobial control program*, Ann Intern Med (1978, 89:793-95); W Schaffner, WA Ray, CF Federspiel, *Surveillance of antibiotic prescribing in office practice*, Ann Intern Med (1978, 89:796-99).

[63] *See:* Schroeder *et al.*, *supra* note 48.

[64] *See:* MW Pozen, H Gloger, *The impact on house officers of educational and administrative interventions in an outpatient department*, Soc Sci Med (1976, 10:491-95); CB Lyle, Jr., DS Citron, WC Sugg, Jr., OD Williams, *Cost of medical care in a practice of internal medicine. A study in a group of seven internists*, Ann Intern Med (1974, 81:1-6).

[65] AR Martin, MA Wolf, LA Thibodeau, V Dzau, E Braunwald, *A trial of two strategies to modify the test-ordering behavior of medical residents*, N Engl J Med (1980, 303:1330-36).

for factors other than the audit-feedback process that could have influenced test-ordering practices. Its weakness, like that of most other studies of such interventions, was its relatively short duration.

Restrictions. A few attempts have been made to restrict various medical resources. Dixon and Laszlo restricted the number of chemistry and hematology tests that interns could order per patient day,[66] resulting in a significant reduction in the number of tests ordered. Although careful studies of the impact of this ceiling on the quality of care were not conducted, the authors apparently felt that there was no evidence that the patients had been deprived of significant benefits (benefits worth their cost).

Another approach that has worked successfully but also has generated considerable resistance is the requirement that a specialist review orders for certain diagnostic tests or treatments before the orders are carried out. The most dramatic example has been a program conducted by McGowan and Finland to control the wasteful utilization of newer, more expensive, and more toxic antibiotics; they achieved a 60-90% reduction in utilization.[67] However, the program was not popular, and utilization increased markedly after the controls were lifted.

Incentives. To some extent, many of the attempts to change resource utilization behavior through education and audit have involved implicit positive incentives: the approval of peers and authority figures. We know of only one published study of the impact of explicit financial rewards to reduce resource utilization, the study by Martin and his colleagues[68] of laboratory test ordering among interns. This study included a group who were given financial incentives: $150-225 worth of free medical journal subscriptions and textbooks. The group provided with such incentives did not changes its laboratory test ordering behavior; indeed, test ordering rose slightly. It is not clear whether this experience was a fair test of the impact of financial incentives in such a setting. It may have been that the relatively small size of the reward was more of an insult than an inducement to the medical residents.

Involving the Doctor in Hospital Cost Control

Prerequisites. Doctors have a dominant role in determining the cost of hospital care, and there are reasons to believe that costs can be contained without compromising the quality of

[66] *See:* Dixon and Laszlo, *supra* note 9.
[67] JE McGowan, Jr., M Finland, *Usage of antibiotics in a general hospital: Effect of requiring justification,* J Infect Dis (1974, 130:165-68).
[68] *See:* Martin *et al., supra* note 65.

care. Therefore it is attractive to consider changes in the relationship of the doctor to the hospital that would encourage cost-minimizing behavior.

If hospital care is to become more cost-effective, two prerequisites must be met: Doctors must not only want to deliver more cost-effective care, but they must also know how to do it. Knowledge without motivation is unlikely to change practice substantially. Motivation without knowledge, is likely to change practice in ways that are not always desirable. Motivation without knowledge is a kind of blunt instrument. For one would rather not discourage all demand for care, only demand that would lead to a waste of resources. And one would rather not encourage a reduction in the use of all diagnostic and therapeutic resources by providers, only a reduction in the wasteful use of resources. In Dyck's successful program to discourage unjustified hysterectomies, for example, the number of justified hysterectomies dropped, as well.[69]

Reinforcing Cost-Effectiveness. We have argued that doctors do respond to incentives to conserve resources. The fact is, however, that the typical doctor now has little reason to be concerned with the cost of hospital care. For one thing, he or she assumes, usually correctly, that the hospital will be reimbursed fairly for services rendered. For another, the doctor assumes no liability for a hospital's operating deficits. In a number of ways, the actions of doctors can introduce inefficiencies into the working of the hospital or can produce frank financial penalties for the hospital but not for the doctor. Examples of such actions include: failure to complete the medical record summary, thereby delaying billing and compromising cash flow; failure to give 24-hour notice of an anticipated discharge—which occurred in 85% of cases studied at one hospital;[70] decisions to hospitalize patients who can be cared for in the ambulatory setting, leading to utilization review denial; inefficient timing of hospitalization (the "Friday exodus and Sunday influx" phenomenon).[71]

Several quite different economic incentives have been proposed for encouraging the doctor to reduce the costs of hospital care. Most often discussed and implemented is the involvement of doctors in prepaid insurance programs (HMOs or IPAs) in which they share the risks of providing costly care. A second proposal is that doctors who order services deemed wasteful accept financial risk along with the hospital, the patient, and the third party,[72] but it seems unlikely that acceptable criteria

[69] *See:* Dyck *et al., supra* note 61.
[70] MC Creditor, *If the doctors owned the hospitals,* N Engl J Med (1971, 284:134-38).
[71] *Ibid.*
[72] JM Eisenberg, AJ Rosoff, *Physician responsibility for the cost of unnecessary medical services,* N Engl J Med (1978, 299:76-80).

for "necessity" could be developed—except for very flagrant (and therefore very unusual) cases.

Creditor has put forward a third approach, arguing that the true consumer is not the patient but the doctor; that, while the hospital benefits from the practice of its medical staff, it also subsidizes that practice—by providing resources (instruments, operating rooms, skilled personnel) upon which the members of the medical staff are completely dependent for their practice, and for which they pay no fee. He proposed that hospitals bill doctors for the services they have used, the way an institution would treat a normal customer.[73] Such an arrangement would have to be tied to an alteration in the way the doctor was reimbursed, of course, but Creditor did not develop this point further.

A fourth interesting idea is that a prospective reimbursement formula for the care of a patient with a particular disease (or diagnostic related grouping) be created that would provide a total sum to be divided by the hospital and the doctor. The doctor would keep a relatively greater fraction of the total sum when relatively fewer hospital resources were utilized, and vice versa. Such an approach could threaten to pit the medical staff against the hospital, but that would not inevitably follow. This idea deserves further study.

An approach of more immediate practicability has been pioneered at the Johns Hopkins Hospital.[74] The hospital has been divided into functional units (cost centers), each of which is headed by a Chief of Service. A detailed accounting of costs and revenues is made for each unit, and each unit is expected to meet budgeted revenue projections (by keeping occupancy rates up) and to control expenses. Since the hospital is reimbursed according to a formula based on diagnostic related group and length of stay, the ordering of resources (other than the use of hospital beds themselves) constitutes an expense, not a source of revenue. The initial experience with this decentralized management system has been favorable.

If doctors are to be involved in the task of containing hospital costs, economic incentives alone are unlikely to be sufficient. The doctors, administrators, and trustees must be better integrated in hospital leadership committees. Indeed, this is already happening. While physician membership on the governing boards of hospitals was once regarded as improper, it has now become common. Conversely, trustee membership on medical staff committees is also a growing practice.[75] It seems likely that the mettle of hospital trustees is likely to be tested as

[73] *See:* Creditor, *supra* note 70.
[74] GD Zuidema, *The problem of cost containment in teaching hospitals: The Johns Hopkins experience*, SURGERY (1980, 87:41-45).
[75] *See:* McNerney, *supra* note 3.

never before in the 1980s. A board of trustees that, in Terrill's words, "like a constitutional monarch . . . reigns but does not rule"[76] will threaten its hospital in the competitive years ahead.

Learning How To Give Cost-Effective Care. We have argued that the doctor often does not know the benefits, risks, and costs of a particular contemplated action, and that lack of such knowledge raises ethical questions when introducing incentives to conserve on resources. Others would counter that, despite the absence of formal studies, doctors know from experience how they can cut corners without compromising the quality of care. We would not deny that for a moment. In support of this argument is the fact that HMOs have succeeded in conserving on hospital care, and that no convincing evidence has been presented that (in established HMOs) care has been compromised in so doing. Nevertheless, we suspect that many medical practices of marginal value are perpetuated out of habit and tradition even in settings where incentives to conserve care are strong. These will not be abandoned unless convincing evidence is presented to challenge their value.

It is surely a pipe dream to expect that pertinent information about benefits, risks, and costs will soon (or ever) be available for most of the different decisions a doctor may be called upon to make. We are fortunate, however, that a relatively few diseases and a relatively few technologies are involved in a large fraction of the decisions made by doctors. By concentrating on learning more about these high-volume problems and technologies, we can have a substantial impact.

The "science of clinical practice" about which we spoke earlier is becoming a reality. Witness the application of probability theory, the development of computerized data bases to which the theory can be applied, the development of cost-effectiveness analyses, the investigation of decisionmaking by doctors and patients. From our keenly biased perspective, we predict that medical historians will regard the growth of this science as one of the important events in medicine during the 1970s. Admittedly, it will not rival the monumental events that should have lifted our depressed spirits over the past decade: the prodigious advances in biology (recombinant DNA technology, the generation of monoclonal antibodies, the surprises of neuroendocrinology), the astonishing achievements in imaging technology (computed tomography, diagnostic ultrasound, positron emission tomography, nuclear magnetic resonance, subtraction angiography), or the new therapeutic armamentarium in cardiology. But it will be noted that in the 1970s the science of clinical practice made its way solidly into the mainstream of the medical literature, entered the argot of the medical

[76] RP Terrill, *Economic relations between hospital and medical staff. The Trustees' role*, TRUSTEE (1970, 23:1-5).

profession, and even crept into the consciousness of a few doctors as they made decisions about the care of patients.

It is fortunate that this has come to pass at a time doctors, hospital administrators, and hospital trustees are grappling with growing pressures to conserve resources. The interests, aspirations, and viewpoints of these three groups are never likely to be entirely consonant. The next few years will remind us all of that fact. Nothing fosters political infighting so much as the absence of pertinent information about the question at hand. Better information about the objective components of clinical decisions—probabilities of benefits and harm, real dollar costs—should help neutralize pressures from any quarter either to waste resources or to withhold necessary resources.

Conclusion

The pressure to use fewer hospital resources raises the same ethical question as does the past pressure to use more of these resources: Does such pressure adversely influence the health of the patient through its effect on the doctor's actions? Given the evidence of waste in the use of current resources, including the evidence of suffering produced by such waste, it is quite possible that the quality of health care would not suffer if, for example, annual health care expenditures were frozen at their current fraction of the gross national product.

But we cannot know this for sure. We cannot know what the ethical consequences may be under any formulas for reimbursement, or with any organizational changes in the relationship of doctors to hospitals, without knowing more about what constitutes "proper medical practice." The definition of "proper medical practice" is often obscure. This is not only because the definition depends in part on value judgments that are inherently variable. It is also because the definition depends on unknown objective information—e.g., the probabilities of various benefits and harms, the real dollar costs of various technologies. It is ethically imperative that we do more to obtain such information. Having done so, will we have solved the ethical dilemma of choosing between a patient's values and those of society, when they are in conflict? No, but we will have allowed this inescapable ethical problem to be posed more clearly when contemplating a particular decision in the care of an individual patient.

I am indebted to Elizabeth Pappius for her invaluable help in surveying the literature and for her editorial comments, to Jean Frier for her review of published data on the decision to hospitalize patients, to Mary Ann Baily, Ph.D., Howard Frazier, M.D., Arnold Epstein, M.D., and Debra Caplan for their critical review of the manuscript, and to Jane Gilligan and Linda McCormick for their help in preparing the manuscript.

The Rationing of Hospital Care

V

E. Richard Brown, Ph.D.*

Rationing, to most Americans, conjures up images of wartime shortages, long waits in service station lines during the oil embargo of the 1970s, and the chronic shortages of consumer goods in many socialist countries. They assume that it is not, and never should be, a normal method of allocating resources in our advanced capitalist society.

In health care, many health professionals assume that rationing is the inevitable and sad result of government interference in medicine, and particularly the consequence of national health insurance, rather than an inherent phenomenon.[1] Others resign themselves to the view that rationing of health services is inevitable in the future due to "spiraling costs and an enfeebled economy," regardless of whether national health insurance is enacted.[2] In the latter view, rationing will become necessary to control the rising costs of medical care, particularly in the hospital sector.[3] Such somber predictions usually carry with them funereal pronouncements of the demise of the public's recently conceded "right to health care" and of the

*School of Public Health, University of California, Los Angeles, California.

October 1981

[1] William R. Felts, *Rationing: The Inevitable Consequence of NHI*, INTERNIST 19(8):22, October 1978.
[2] Edmund D. Pellegrino, *Rationing Medical Care: The Need for Distributive Justice*, INTERNIST 19(8):5, October 1978.
[3] Walter J. McNerney, *Allocating Resources: Who Will Decide?* INTERNIST 19(8):6, 8, October 1978; James H. Sammons, *Rationing Medical Care: Problems Worldwide with Resources vs. Needs*, INTERNIST 19(8):21, October 1978; John P. Mamana, *Rationing Health Care Services Through Regulation*, HOSPITAL MEDICAL STAFF 8(3):7-9, March 1979; and *Rationing of Care Seen as Inevitable Cost Control Tool*, HOSPITALS 53(20):19, October 16, 1979.

physician's ability to order whatever care is deemed necessary for his or her patient.

In reality, medical care has always been rationed—that is, its allocation or distribution process has been controlled—by the ability to pay for care, by the supply and distribution of medical resources and personnel, and by the way health professionals allocate their efforts.[4] In the United States these rationing mechanisms have been attributes of the market system in which health care is produced and distributed. In a market system each unit (e.g., each hospital, physician, drug company) is free to produce and distribute its goods or services within the constraints of market conditions. These conditions include the availability and costs of inputs (labor, materials, and capital), the supply and prices of competing goods or services, and the economic demand for the goods or services. In medicine, these basic market conditions have been modified by licensing and other forms of regulation, but individual units simply take these modifications into consideration as additional market constraints.

Each unit thus decides how it will allocate its medical resources within the conditions set by its particular market. For example, the medical staff, administrator, and trustees of a hospital may decide to follow their traditional patient population from their former inner-city community to the suburbs, abandoning the old community's newer (black or Hispanic) population. Similarly, the medical staff of a hospital that serves primarily middle- and upper-class white communities may grant staff privileges only to those physicians whose medical practice and patients "fit in." Finally, many private hospitals have for years been transferring poor or uninsured patients to public hospitals, admitting only those persons who are well insured or are affluent enough to pay the high cost of hospital care. The widespread assumption that rationing is something imposed by government on medicine is testimony to how natural the market system seems to most of its practitioners—and thus how invisible its mechanisms have become. But the market system is no more "natural" than any other allocation system.

First we will examine the development of hospitals in the United States, the changing forms rationing has taken, and the effects of the different forms. Some of these allocation methods—such as private health insurance, the Federal government's Hill-Burton program, and the Medicare and Medicaid programs—have redistributed hospital care more equitably between urban and rural areas and between social classes. Some governmental rationing methods—such as some cost-contain-

[4] David Mechanic, *Rationing of Medical Care and the Preservation of Clinical Judgment*, JOURNAL OF FAMILY PRACTICE 11:(3):431-433, September 1980.

ment measures—have actually increased the gaps in access to hospital care between the poor and the nonpoor. But both outcomes have resulted from efforts to reduce problems generated by the market autonomy of hospital administrators and the professional autonomy of physicians. Public control was necessitated by the rapidly rising costs of hospital and other medical care that accompanied the introduction of massive government subsidies into the privately controlled medical marketplace. These subsidies were themselves created as an antidote to the tendencies inherent in the market system to develop, produce, and distribute health care according to criteria other than the health needs of the population. In other words, subsidies were introduced to allocate hospital and other health care more equitably—to distribute care on the basis of health need, rather than according to social class, race, or location. However, the public's open-ended fiscal involvement was not well served by hospitals' market autonomy, necessitating Federal and state action to protect the public investment. Finally, we will examine the forms these policies took and their consequences.

The Changing Allocation of Hospital Care

The Origins of American Hospitals. The general hospital of today grew out of public and private institutions that, until the twentieth century, were "largely undifferentiated welfare institutions" for only the poorest classes in society.[5] Some of the earliest public hospitals originated as workhouses or almshouses, providing custodial care for paupers, the physically and mentally disabled, and the terminally ill. By the nineteenth century, these poorhouses began developing special facilities for the sick within their walls, but public hospitals have never completely shed the reputation associated with their birth.

Private hospitals were also formed exclusively to serve the poor. While public hospitals, in the tradition of local governmental responsibility for the poor, were supported by local taxation, private hospitals were charitable institutions reflecting the benevolent paternalism of their affluent donors and trustees—who used means tests and other devices to assure that only the "deserving" poor would use their services and to help suppress "idleness and begging."[6]

The stigma associated with hospitals usually assured that only the poor would seek admission. The middle and upper

[5] Morris J. Vogel, *The Transformation of the American Hospital, 1850-1920*, in HEALTH CARE IN AMERICA: ESSAYS IN SOCIAL HISTORY, S. Reverby and D. Rosner, eds. (Philadelphia: Temple University Press, 1979), 105.

[6] *Ibid.*, 105-116. On charity in the nineteenth century, see Robert H. Bremner, AMERICAN PHILANTHROPY (Chicago: University of Chicago Press, 1960).

classes, and even those strata of the working class who could afford the luxury, were treated by physicians in their offices or, more usually, in the patient's home. Thus, up to the end of the nineteenth century, hospital care was rationed exclusively to the poor.

Hospitals for the Affluent, Hospitals for the Poor. By the end of the nineteenth century, private hospitals were becoming increasingly differentiated from public hospitals. Private institutions developed new roles and clients in a growing medical market, while public hospitals were left to serve the groups and roles excluded from that market. Three developments encouraged this dramatic change.

First, discoveries and inventions resulting from medical research introduced new technological developments for diagnosis and treatment that raised the cost of hospitalization. The use of oxygen, rubber gloves, frequently changed dressings, and aseptic conditions for surgery were only a few of the innovations that increased the cost of hospital care.

Second, while the costs of patient care were rising, the depression of the 1890s increased the number of destitute patients needing hospital care; but the same economic conditions dampened the generosity of donors to private hospitals. Medical costs and patient loads thus rose while hospital income declined. In the resulting fiscal crisis, many private hospital trustees reconsidered their hospitals' traditional role of caring only for the poor.

Both of these developments provided a common bond between hospitals and physicians, whose changing medical practice was the third influence on private hospitals. The new medical technologies forced physicians to look for new sources of capital, facilities to house and personnel to operate the new technologies, and nursing staff to care for their patients. The hospital seemed the logical repository for expensive medical technologies and gradually emerged as the "doctor's workshop."

In order to attract middle- and upper-class paying patients and thereby reduce or eliminate their financial problems, many private hospitals built new facilities or renovated old ones, separating paying patients' private accommodations from the charity wards. Brooklyn Hospital, as an example, was able to increase the proportion of paying patients from 16% of all patient days in 1895 to 45% in 1905. With this momentous change in private hospitals, "the provision of hospital care ceased to be an act of charity," David Rosner notes, "and became a commodity to be bought. . .by those who could afford it."[7]

[7] David Rosner, *Business at the Bedside: Health Care in Brooklyn, 1890-1915*, in HEALTH CARE IN AMERICA, 118.

Thus, hospital care in the United States was transformed into two relatively distinct sectors. Private hospitals, both voluntary (also called charitable or not-for-profit) and proprietary (or for-profit) institutions, developed a fairly lucrative market among the upper working class, the middle class, and the wealthy; they saw charity care for the poor as an inconvenience at best. Public hospitals, and a shrinking number of voluntary ones, were left to care for the working and nonworking poor. As we will see, each subsequent major development in the financing of hospital care reinforced this separation—even as more of the population gained access to private hospitals.

Reinforcing the Rationing of Hospital Care. Although the poor were priced out of private hospital care, there was growing concern among physicians and private hospitals that individuals who could afford to pay for their own care were taking advantage of reduced-fee and free care and thereby depriving private providers of their rightful fees. Charity hospitals and clinics (which generally considered the poor "clinical material" for associated medical education programs) and public medical services were attacked by local medical societies for failing to enforce strict means tests on their patients. These attacks had their intended effects of narrowing further the ranks of those permitted to obtain free or reduced-fee hospital or other medical care at the same time that the poor were effectively being excluded from mainstream private practice.

If spartan facilities, bureaucratic organization, less-than-benign paternalism of many medical and nonmedical staff members, and stigmatizing means tests were not enough to restrict public hospitals to the poor, many states added legal restrictions to prevent any but the indigent from using public facilities. In 1936, a California appellate court, for example, ruled that the state's county hospitals could treat only persons unable to afford the full cost of private care, accepting indigent persons as nonpaying patients and charging persons judged able to pay part of the cost according to their means. The court prevented county hospitals from competing with private hospitals for full-paying patients on the grounds that local governments may not "engage in private business or enterprise" or use "public money for private purposes."[8]

Clearly, by the time of the Great Depression both public and private hospital care were being rationed. Private hospital care was generally available only to persons who could pay their own way; that is, it was rationed by price. Some private hospital

[8] Goodall v. Brite, 11 Cal. App. 2d. (1936) 541, 544. This prohibition against public hospitals serving the general community continued until the implementation of the Medicaid program necessitated a change. In California, the Medi-Cal implementation legislation, passed by the state legislature in 1965, authorized county hospitals to serve persons irrespective of their ability to pay.

charity care was provided to the poor, largely for teaching purposes and mainly in wards separated from paying patients. Both public hospital care and private charity care were rationed to the poor on the basis of strict means tests.

Because the costs of private hospital care made it less available to the poor and because of the degradation that accompanied charity and public hospital care, persons of lesser means received substantially less medical care than more affluent groups. In a study conducted between 1928 and 1931, the Committee on the Costs of Medical Care found that persons in upper-income families averaged one and one-half times as many hospital admissions each year as members of low- and moderate-income families. In fact, hospital care and other types of medical care varied directly with income.[9]

The commodity nature of hospital care meant that its distribution paralleled the class structure of society. In addition to the inequalities in the use of hospital care produced by this market distribution, many working-class and middle-class people who could afford ordinary outpatient medical expenses were either impoverished by the extraordinary expenses of hospitalization or were forced to forego such care when it was needed.

The rationing of hospital care was not only unsatisfactory for a large portion of the American public, however. It also created severe financial instability for hospitals, particularly during frequent downward business cycles. As the Great Depression deepened, many hospitals tottered on the edge of insolvency and some were forced to close their doors. Inequities in financial access, impoverishment due to large hospital bills, and unstable revenues jointly formed the base for a partial solution—the growth of hospitalization insurance.

Socializing the Risks and Payment of Hospital Care. To alleviate the deteriorating financial condition of the hospital industry, the American Hospital Association actively expanded its Blue Cross hospital insurance plans throughout the country. First introduced in 1929 by Baylor University Hospital in Dallas, Texas, the plans grew slowly but proved their usefulness and popularity. By 1940, 12 million Americans (9% of the population) were enrolled in hospital insurance plans, half of them in Blue Cross. About 5 million (4% of the population) also were covered for surgical benefits.[10]

---

[9] I.S. Falk, C. Rufus Rorem, and Martha D. Ring, THE COSTS OF MEDICAL CARE, Committee on the Costs of Medical Care Publication No. 27 (Chicago: University of Chicago Press, 1933), 599.

[10] Marjorie Smith Mueller, *Private Health Insurance in 1973: A Review of Coverage, Enrollment, and Financial Experience*, SOCIAL SECURITY BULLETIN 38(2):21-40, February 1975; and Marjorie Smith Mueller, *Private Health Insurance in 1975: Coverage, Enrollment, and Financial Experience*, SOCIAL SECURITY BULLETIN 40(6):3-21, June 1977.

For those individuals and families able to buy health insurance, the financial risks of hospitalization were socialized—that is, spread among many persons, enabling them to obtain expensive medical care without necessarily being impoverished. Hospitals benefited from the socializing of risks and payments for hospital care because their insured patients represented a growing market and an increasingly stable source of revenues. These benefits and the pressure by labor unions for nontaxable fringe benefits encouraged the spread of work-related group plans. In 1945, after several years of cost-plus government war contracts and union pressure on employers, enrollment in hospital insurance plans reached 32 million persons, nearly one in every four persons in the country. After the war, commercial insurance companies aggressively entered the health insurance market they had previously all but ignored. By 1950, 77 million people, half the population, were covered by private hospital insurance plans; and by 1960, more than 130 million people were enrolled.[11]

But private health insurance had its limitations. First, private insurance does not cover all services equally well. Because of the greater expense of hospitalization and the efforts of the hospital industry to develop and promote coverage for its services, hospitalization insurance has always been more widespread than other health insurance benefits. In 1950, for example, only two-thirds of the number of persons enrolled in hospitalization insurance plans had surgical benefits. By 1960, 91% of those covered for hospital care were insured for surgical benefits, but even by 1975, the proportion had increased only an additional 1%.[12] And an even smaller percentage of the population in each period has been insured for nonsurgical physicians' services, prescription drugs, or dental care.

Second, private insurance does not meet all the costs associated with covered services. In order to reduce utilization and costs, health insurance plans usually restrict their coverage with exclusions (medical conditions and services not covered), deductibles (out-of-pocket payments required before benefits begin), copayments (a percentage of the costs beyond the deductible for which the patient is responsible), and limitations (upper dollar limits for an illness or year beyond which the plan's coverage ends). These rationing devices have substantially limited the financial benefits of covered services. For example, in 1950, when 51% of the population was enrolled in hospital insurance plans, these plans met only 37% of insured patients' expenditures for hospital care (even less if the costs of insurance premiums are added to patients' expenditures). Because of increasingly complete hospitalization insurance included in work-related group plans, hospital care is now more

[11] *Ibid.*; Mueller, *Private Health Insurance in 1973*, and . . . *in 1975*.
[12] Calculated from Mueller, *Private Health Insurance in 1975*.

fully covered. In 1975, private hospitalization insurance covered 80% of the population and met 80% of patients' hospital expenditures.[13] But physicians' services and other types of medical care, particularly preventive health services, are less well covered.

The third limitation, and perhaps the one that most seriously affects the allocation of hospital care, is that not all population groups are equally well covered. The most comprehensive health insurance has been and still is available mainly to the upper working class and to the more affluent portion of the population—that is, persons covered by work-related plans and those who can afford individual insurance policies. In 1958, 86% of the upper-income third of all American families had some kind of private health insurance, while only 42% of the lower-income third had any coverage.[14]

Like the lower working class and the poor, the elderly have been unable to obtain adequate private hospitalization coverage at a price they can afford. In 1962, four years before the introduction of Medicare, approximately half of all persons aged 65 and over had no insurance for hospital care, compared with only about one in four persons under 65.[15] The elderly population has a substantially greater need for hospital care than the rest of the population. In 1965, before Medicare and Medicaid and with less than adequate insurance available to them, persons aged 65 and over averaged nearly three times as many days of care in short-stay hospitals as all age groups.[16] The elderly's greater need for such care made them bad insurance risks, pricing many of them out of the insurance market and forcing them to postpone or forego care and to become impoverished when hospitalization was unavoidable.

Health insurance thus represented a step forward in bringing the benefits of more expensive kinds of medical care to more people by socializing the risks and payment for that care. However, to control costs and assure profitability, the insurance industry:

(1) marketed their services mainly to the upper working class and the more affluent population, and therefore failed to meet the needs of the low-paid, nonunionized working population or the elderly;

---

[13] *Ibid.*

[14] Herman W. Somers and Anne R. Somers, DOCTORS, PATIENTS AND HEALTH INSURANCE (Washington, D.C.: The Brookings Institution, 1961), 367.

[15] Mueller, *Private Health Insurance in 1973.*

[16] Cambridge Research Institute, TRENDS AFFECTING THE U.S. HEALTH CARE SYSTEM (Washington, D.C.: Department of Health, Education, and Welfare, January 1976), 11.

(2) insured the somewhat less elastic types of care, favoring hospital care over ambulatory services and curative over preventive care; and

(3) rationed care by several cost-sharing methods, requiring the insured to weigh the out-of-pocket costs of care against the felt urgency of the medical need. Each of these allocative mechanisms favors higher-income population groups who are better able to afford the added costs of providing themselves with comprehensive medical care.

Subsidizing the Supply of Hospital Beds: Public Capital. In addition to the problems of paying for care, one of the most significant inequities in the allocation of hospital care has been the supply and distribution of hospital beds. In 1928, the Committee on the Costs of Medical Care found that the Southern states averaged about 2 community hospital beds per 1000 population compared with approximately 5 beds per 1000 in the North Atlantic and Western states. Some states, like Wisconsin with 6.5 beds per 1000 population, were even by today's standards rich in hospital resources; while others, like South Carolina with only 1.3 beds per 1000 population, were resource-poor. Within states, rural areas and towns lacked the hospital resources available in urban areas. Although jurisdictions smaller than 10,000 persons included 58% of the nation's population in 1928, they accounted for only 42% of the nation's general and specialty hospital beds. On the other hand, cities of 10,000 to 100,000 persons represented only 16% of the population but 28% of all hospital beds; and cities larger than 100,000 persons represented 26% of the population and 30% of all hospital beds. Many rural counties—43% of all counties—had no community hospital at all. The black population, particularly in the South where segregation of hospitals was rigidly practiced, suffered from a "scarcity of hospitals."[17]

The inability or unwillingness of many local private and public sources of capital to finance hospitals in their communities necessitated Federal intervention. After heavy lobbying by the American Hospital Association, the American Medical Association (AMA), and the Truman administration, the Congress passed the Hospital Survey and Construction Act of 1946, better known as the Hill-Burton Act, the major vehicle for subsidizing the nation's supply of hospital beds. From its inception in 1947 through 1974, the Hill-Burton program spent more than $4 billion in Federal funds on grants and loans for construction and modernization of medical facilities. The program also generated an additional $10.4 billion in state and local matching funds for the same projects. Hill-Burton grants

[17] Allon Peebles, A Survey of Statistical Data on Medical Facilities in the United States (Washington, D.C.: Committee on the Costs of Medical Care, 1929), 38, 40, 91-93.

and loans have played a role in the construction of nearly 40% of all beds in 6000 of the nation's not-for-profit and public community hospitals.[18]

The Hill-Burton program had a substantial impact on both the total supply and the distribution of hospital beds. In 1946, the eight states with the highest per capita income had 4.0 community hospital beds per 1000 population, while the poorest eight states had only 2.5 beds per 1000 population. By 1971, those eight wealthiest states had 4.1 beds per 1000 population, but the eight poorest states had eliminated their disadvantage, surpassing the affluent states with 4.3 beds per 1000 population.

The development of many small, rural public and private hospitals funded by Hill-Burton substantially reduced the rural disadvantage.[19] However, despite amendments in 1964 and 1970 intended to encourage modernization of urban hospitals and construction of hospital facilities in poverty areas, the program was not successful in equalizing the availability or quality of hospital services in urban and rural poverty areas. There are two reasons for this failure. First, the Hill-Burton matching-grant requirement and the loan and loan-guarantee program have proved difficult for many poor communities. Many of them have been unable to raise their local share or get credit from lending institutions. Others have lacked sufficient political clout to persuade local or state governments to fund public facilities in their areas. In Chicago, for example, up to 1973 not one inner-city community hospital had been helped by the Hill-Burton program, although two dozen outlying and suburban hospitals had been helped to expand or build new facilities.[20]

Second, many private, not-for-profit hospitals in inner-city neighborhoods—including many hospitals aided by Hill-Burton—have abandoned their service areas either by transferring their facilities to outlying suburban areas or by establishing satellite facilities in suburban areas that then drain resources away from the already resource-poor inner-city communities. In a study of 18 cities in the Northeast and Midwest, Alan Sager found that in neighborhoods that were three-fourths or more black in 1970, 47% of the hospitals that existed in 1937 had closed or relocated by 1977, compared with only 14% in neigh-

[18] Cambridge Research Institute, TRENDS, 92; and U.S. Senate, Labor and Public Welfare Committee, *Legislative History, P.L. 93-641, Senate Report 93-1285*, November 12, 1974 (St. Paul, Minn.: West Publishing Co., 1975), 7862.

[19] *Ibid.*, 7863 and 7881; and Judith R. Lave and Lester B. Lave, THE HOSPITAL CONSTRUCTION ACT: AN EVALUATION OF THE HILL-BURTON PROGRAM, 1948-1973 (Washington, D.C.: American Enterprise Institute, 1974), 27, 38-39.

[20] Cambridge Research Institute, TRENDS, 94.

borhoods that were one-fourth or less black in 1970.[21] Those that remain in poor communities are deteriorating. By the mid-1970s, more than a third of the country's hospital beds—and particularly those in poverty areas—were in facilities needing remodeling or replacement.

In addition to improving the supply and distribution of hospital beds, the Hill-Burton legislation attempted to insure that the use of public funds to subsidize capital investment in hospitals did not benefit only those persons who could afford to pay for care. The legislation required hospitals that received Hill-Burton grants or loans to provide "community service" and "charity care" (or uncompensated care) under one of several different formulas. However, implementation by both the Federal government, and the state Hill-Burton agencies has been lax or nonexistent, "a sorry performance" in the words of the Senate Committee on Labor and Public Welfare.[22] Despite new laws and regulations designed to strengthen Federal enforcement of these Hill-Burton provisions, compliance by hospitals remains spotty; and numerous instances have been documented of Hill-Burton hospitals denying even urgent care to indigent patients.[23]

Thus, the injection of public capital into the hospital sector succeeded in expanding the supply of hospital beds and substantially redistributing that supply among the states and between urban and rural areas. But the program's dependence on state and local initiative and the raising of state and local matching funds favored communities with sufficient capital or political resources. Economically underdeveloped or depressed areas have been unable to obtain their fair share of these public funds. Furthermore, the market autonomy of the private hospital sector together with political power, particularly at the state level, have thwarted the program's intention of using the fruits of public capital to improve the access of the poor to private hospital care. Although the availability of hospital care was improved by increasing the total supply of beds and reallocating their geographic distribution, the infusion of public capital to the hospital sector did not solve the continuing problem of inequity for all income groups.

---

[21] Alan Sager, *Urban Hospital Closings in the Face of Racial Change*, Health Law Project Library Bulletin 5(6):169-181, June 1980; Valerie A. Seiling, *The Prima Facie Case and Remedies in Title VI Hospital Relocation Cases*, Cornell Law Review 65(4):689-714, April 1980; and U.S. Senate, Labor and Human Resources Committee, *Legislative History, P.L. 96-79, Senate Report 96-96*, April 26, 1979 (St. Paul, Minn.: West Publishing Co., 1980), 1396.

[22] U.S. Senate Labor and Public Welfare Committee, *Legislative History, P.L. 93-641*, 7900.

[23] Armin Freifeld, *Hill-Burton: The Shape of Things Today and Yet to Come*, National Health Law Program, 1981.

Public Subsidy for the Hospital Care Market. To achieve equity for all in medical care, proponents of comprehensive national health insurance doggedly submitted bills in the Congress during the 1930s and '40s, only to face certain defeat at the hands of the AMA's well-funded lobbying machine and national campaigns. By 1950, "realism" dictated a retreat. The spread of private health insurance, especially to workers in unionized industries, had undermined the labor movement's political support for national health insurance. The liberal political programs that had a broad base of popular support during the 1930s and '40s were under attack in the early 1950s, as McCarthyism and the Korean War dominated the political landscape.

President Truman's chief Social Security advisers proposed backtracking from a comprehensive insurance plan for everyone in the country to a set of limited hospital benefits restricted to beneficiaries of the Social Security pension program. Truman's advisers believed that by backing away from a universal and comprehensive plan, excluding physicians' services, and limiting the program to Social Security beneficiaries, they could side-step opposition from the AMA and at the same time win broad support.

Despite these retreats, little legislative progress was made until 1960. The ever increasing costs of medical, and especially hospital, care and the increasing clamor from senior citizens' groups and labor made it apparent that some Federal program would have to be enacted to help the elderly meet the high cost of getting sick. After Representative Wilbur Mills's House Ways and Means Committee killed a proposal for a Social Security-linked hospital insurance plan, Mills and Senator Robert Kerr introduced a bill to expand the existing system of Federal grants to the states for "vendor payments" (payments made by local welfare agencies to hospitals and doctors for medical care provided to public assistance recipients). The Kerr-Mills Act covered medical care for the elderly, including those receiving cash assistance and those too poor to pay their own medical bills even though ineligible for cash grants (the latter group were designated "medically needy").[24]

Although the states were slow to join the program, 44 of the 54 states and territories had some program in effect by 1965. However, 5 large and relatively affluent states—New York, California, Massachusetts, Minnesota, and Pennsylvania—which together included 31% of the elderly population in the country, were receiving 62% of the Federal funds for the medically

[24] Robert Stevens and Rosemary Stevens, WELFARE MEDICINE IN AMERICA: A CASE STUDY OF MEDICAID (New York: The Free Press, 1974), 26-30; and Theodore R. Marmor, THE POLITICS OF MEDICARE (Chicago: Aldine Publishing Co., 1973).

needy program. The Kerr-Mills program substantially increased the flow of public funds to hospitals and nursing homes, which together accounted for two-thirds of total vendor payments of $1.4 billion in fiscal year 1965-66.[25]

The Kerr-Mills Act was unsatisfactory to just about all concerned parties—particularly to senior citizens' groups, which objected to the indignities of a means test in the welfare-linked program, and to the states, which wanted relief from the rising costs of welfare medical care. By 1965, pressures from senior citizens' groups, labor unions, and civil rights organizations, reform-minded health professionals, and their supporters in the Democratic administration and the Congress were able to overcome the AMA's powerful opposition to any form of government-run health insurance.

The results of this coalescence of forces were the Medicare and Medicaid programs, Titles 18 and 19, respectively, of the Social Security Act. Medicare is a federally run insurance program, with Hospital Insurance (HI) financed by a compulsory Social Security tax paid into a special trust fund, and a voluntary Supplementary Medical Insurance (SMI) portion (covering mainly physicians' services) financed by insurance premiums and matching Federal general tax revenues. By 1978, enrollment in Medicare's HI and SMI programs both topped 26 million persons, covering nearly the entire elderly population in the U.S., although with substantial deductibles and cost sharing adopted from private insurance plans.[26]

Medicaid is a Federal-state program, with Federal shares to the state-run programs inversely related to state per capita income. Medicaid enrollment reached 23 million in 1976, but since then it has been gradually declining due to tightened eligibility.[27] Because of the minimal program requirements imposed on the states and the flexibility granted them in setting income eligibility, somewhere between 40% and 67% of all persons in the U.S. with incomes below the poverty line are not eligible for Medicaid.[28]

Impact of Third-Party Programs on the Use of Health Services. Private health insurance, Medicare and Medicaid, and expanded public hospital and clinic services have substantially reduced the inequities in the use of health services noted by the Committee on the Costs of Medical Care a half-century ago.

---

[25] Stevens and Stevens, WELFARE MEDICINE IN AMERICA, 31-36.
[26] *Medicare: Persons Enrolled, 1979*, HEALTH CARE FINANCING NOTES (Baltimore: Health Care Financing Administration, January 1981), 3.
[27] Medicaid/Medicare Management Institute, DATA ON THE MEDICAID PROGRAM: ELIGIBILITY, SERVICES, EXPENDITURES, 1979 Edition (Revised), (Baltimore: Health Care Financing Administration, 1979), 33.
[28] *Ibid.*, pp. 62-63; and Karen Davis, Marsha Gold, and Diane Makuc, *Access to Health Care for the Poor: Does the Gap Remain?* ANNUAL REVIEW OF PUBLIC HEALTH 2:159-182, 1981.

Whereas the Committee's studies found that upper-income groups had far more physician visits and hospital admissions than low-income groups, recent National Health Interview surveys found that poor adults actually make more doctor visits and poor persons of all ages average more hospitalizations than the nonpoor.

Despite these strides, some important deficiencies remain. Children in families below the poverty line average fewer physician visits than nonpoor children, and the poor of all ages receive less dental care and less preventive medical care than persons above the poverty line. Furthermore, when people are grouped according to their health status, middle- and upper-income groups use health services more than do the poor in comparable health conditions. For example, among persons grouped according to number of days of bed disability, upper-income persons make three-fourths more physician visits each year than lower-income persons.[29]

Finally, low-income persons are less likely than more affluent persons to have a regular source of medical care, and low-income persons with a regular source are much more likely to rely on hospitals and clinics as that source.[30] One reason for this difference in sources of care is that most physicians in private practice treat few, if any, Medicaid patients. Indeed, 20% of all physicians see no Medicaid patients at all in their practices, and just 6% of all doctors care for one-third of all Medicaid patients.[31]

Most of these inequities in source and amount of ambulatory care may seem unrelated to hospital care, which is inversely related to income both for rates of hospitalization and for number of days of care.[32] However, it is physicians who admit patients to hospitals; so inequities in sources of physician care and in amount of care are likely to contribute to inequities in hospital care. Thus, the relatively greater use of hospital services by the poor may still not be commensurate with their medical needs.

It is also clear that Medicaid patients and inadequately insured and uninsured low-income patients tend to get channeled to public hospitals and to a relatively few voluntary (not-

[29] Davis *et al.*, *Access to Health Care for the Poor*; and HEALTH OF THE DISADVANTAGED (Hyattsville, Md.: Department of Health and Human Services, 1980), 59, 61.

[30] HEALTH OF THE DISADVANTAGED, 68, 70; Davis *et al.*, *Access to Health Care for the Poor*; Margaret C. Olendzki, Richard P. Grann, and Charles H. Goodrich, *The Impact of Medicaid on Private Care for the Urban Poor*, MEDICAL CARE, 10(3):201-206, May-June 1972.

[31] Janet Mitchell and Jerry Cromwell, *Medicaid Mills: Fact or Fiction?* HEALTH CARE FINANCING REVIEW 2(1):37-49, Summer 1980.

[32] THE NATION'S USE OF HEALTH RESOURCES, 1979 (Washington, D.C.: National Center for Health Statistics, 1980), 58-61.

for-profit) hospitals. Although supporting evidence for this pattern is mostly anecdotal or case material related to lawsuits on behalf of the poor, there is also illustrative data from several localities. In Oakland, California, for example, the county hospital serving the northern part of Alameda County represented 8% of all hospital admissions in the county in 1974, but it accounted for 17% of all Medicaid admissions in the county.[33] The widespread practice of private hospitals and physicians "dumping" uninsured patients and those considered undesirable onto public hospitals has been well documented. In 1979, for example, 21,000 such patients were transferred to Los Angeles County hospitals by private hospitals.[34]

Thus, despite the substantial improvements that have been made in equalizing the volume of hospital care among poor and nonpoor groups, it is unlikely that even the present rates fully match the demonstrably greater needs of the poor. And, it is evident that the poor, particularly the uninsured and inadequately insured working poor, remain at least partially segregated in public hospitals and a small number of charity-oriented private hospitals.

The Gilded Age of Hospitals. The socializing of payment through private hospital insurance, the injection of public funds as a direct subsidy to hospital capital, and the infusion of public subsidies for low-income and elderly patients into the marketplace for hospital care have together transformed the hospital industry from a simple and marginal economic sector to a vast, complex, and profitable one. The period from 1950 through the early 1970s represented the gilded age of hospitals.

In the first 15 years of this period, private health insurance and the Hill-Burton program nearly tripled the resources flowing to the hospital sector. As national health expenditures climbed from $13 billion and 4.4% of the Gross National Product (GNP) in 1950 to $42 billion and 6.1% of the GNP in 1965, hospitals increased their share from $4 billion (30% of the total) to $14 billion (33% of the total). In the next 10 years, however, hospitals absorbed national resources at an even faster rate of growth. National health expenditures grew to $132 billion in 1975, an impressive 8.6% of the GNP. Hospitals increased their share to $52 billion (40% of total health expenditures) and 3.4% of the entire GNP (compared to 1.3% of the GNP spent on

[33] Alameda County Hospitals Project, Final Report (Oakland, Calif.: Alameda County Health Care Services Agency, June 1978), Vol. 2, part a, 14.

[34] Geraldine Dallek, *America Severs Decent Health Care for Millions of Poor*, Los Angeles Times, August 10, 1980, part 5, 3; Milton I. Roemer and Jorge A. Mera, *"Patient-Dumping" and Other Voluntary Agency Contributions to Public Agency Problems*, Medical Care 11(1):30-39, January-February 1973; Barry Roth, *Patient Dumping*, Health/PAC Bulletin, No. 58:6-10, May-June 1974.

hospital care in 1950).[35] Hospitals had become big business, indeed.

The earlier growth in hospital expenditures was clearly fed by the expansion of private health insurance. As enrollment in private hospital insurance plans increased from half the population in 1950 to three-fourths in 1965, the proportion of total national expenditures for hospital care met by private health insurance payments increased from 18% to 42%. With the launching of the Medicare and Medicaid programs, however, total government payments for hospital care jumped from 39% of total hospital costs in 1965 to 55% in 1967 and have remained at or near that level ever since, while the private health insurance share has gradually declined to about 35%. By 1970, Medicare and Medicaid, which did not even exist in 1965, were themselves contributing more than 1 out of every 4 dollars received by hospitals for patient care. In 1979, hospitals collected $30 billion, 35% of their total revenues, from Medicare and Medicaid.[36]

With a steadily growing source of patient care revenues and assurances that Blue Cross and other insurers would reimburse them for interest and principal, hospitals borrowed heavily to finance new construction, modernization, the purchase of new medical technologies, and expansion in the number of their beds. The Hill-Burton program provided a strong stimulus to hospital construction in the late 1940s and 1950s,[37] but its importance was gradually overshadowed by debt-financing as hospital revenues grew without restrictions and provided an incentive for borrowing against future income. Between 1950 and 1965, hospitals increased their assets from $8,612 per bed to $22,084 per bed, an average increase of 10.4% per year. In the next ten years, however (on what seemed the firm foundation of Medicare and Medicaid, as well as the continued growth of private insurance), hospital assets per bed grew an average of 12.6% each year, reaching nearly $50,000 in 1975.[38]

With confidence in their seemingly irrepressible markets, hospitals turned from government and philanthropy as major sources of capital to their own borrowing resources. Debt financing of construction increased from 40% of construction capital in 1968 to 68% in 1976, while government funds fell from 23% of construction capital to 11% in this same period, and philanthropy from 21% to 8%. The decline in government grants did not mean that hospitals were financing their construction,

---

[35] Robert Gibson, *National Health Expenditures, 1979*, HEALTH CARE FINANCING REVIEW 2(1):1-36, Summer 1980.

[36] *Ibid.*; Mueller, *Private Health Insurance in 1973*, and . . . *in 1975*.

[37] See U.S. Senate Labor and Public Welfare Committee, *Legislative History, P.L. 93-641*, 7863.

[38] HOSPITAL STATISTICS, 1977 Edition (Chicago: American Hospital Association, 1977), 4.

expansion, and modernization without government assistance. In the latter 1970s, state and local government-issued tax-exempt bonds were the largest single source of capital for hospital construction,[39] providing hospitals with low-interest loans and permitting greater expansion of assets than would otherwise be possible.

However, this market-generated construction and expansion has contributed to what is generally regarded as an excess of hospital beds and an inappropriate emphasis on expensive types of medical care. With rapidly rising hospital costs came growing pressure for restraining those increases. Blue Cross subscribers in several states organized opposition to the continuous and large increases in premiums. Blue Cross plans were forced to revamp their cozy and nepotistic relationships with the hospital industry, gaining enough independence to begin to look critically at the costs for which they (and ultimately their subscribers) were reimbursing hospitals. The Federal government and many states similarly began exerting pressures through Medicare and Medicaid reimbursement programs and through a series of efforts to directly regulate hospital expansion and bed supply. Upper pressure of the watchful eyes of third-party payers, the average length of stay in hospitals began to decline in 1969, despite the continuing increase in the bed supply (from 3.26 beds per 1000 population in 1950 to 3.75 beds per 1000 in 1965 to 4.35 beds per 1000 in 1975. As the bed supply continued to increase and length of stay continued to decline, occupancy rates also began to fall from their peak of 79% in 1969 (81% for not-for-profit hospitals).[40] With declining occupancy rates, the costs per patient day further increased—at the very time when cost-containment efforts and unfavorable general economic conditions were beginning to tighten the screws on hospital finances, putting as many as one-fourth of all community hospitals in the red and encouraging efforts to return the country (and each area within it) to a ratio of 4.0 beds (or fewer) per 1000 population.

In addition to an expansion in the supply of hospital beds, the increased flow of public and private third-party payments accelerated what became a virtual medical technology race among hospitals, particularly those in the private sector that had the patient-care revenues with which to finance their acquisitions. Market competition among hospitals to attract physicians and their well-insured patients, supported by institutional and community pride, encouraged hospital administrators and boards of trustees to provide the most advanced

[39] Capital Formation in Health Care Facilities (Washington: Health Resources Administration, 1979), 10.
[40] Hospital Statistics, 1980 Edition (Chicago: American Hospital Association, 1980), 5-6.

medical technology and the most modern facilities and services that money could buy.

Physicians, of course, play a central role in generating the demand for hospital care as well as for their own services. Physicians' training in high-technology teaching hospitals, the threat of malpractice suits, and especially the financial incentives in the fee-for-service system all encourage the "technological imperative"—the attitude that if something *can* be done for a patient, it *should* be done.[41] The ordering of excessive diagnostic and therapeutic procedures is supported not only by these hospital and physician characteristics. It is also encouraged by medical technology companies that promote their products to as many hospitals and doctors as they can—without having to prove the effectiveness or safety of their wares and without concern for medical need.

This pattern of rapid, unplanned expansion of advanced medical technologies undoubtedly has provided some benefits to patients with life-threatening illnesses, but the medical effectiveness of only a very few technologies has been adequately assessed. Virtually none has been evaluated in the context of health care delivery systems, considering if other health care investments might provide a greater benefit for more people. For example, would an additional open-heart surgery unit in a community be of more value, equal value, or less value to the community's health needs than a similar investment in, say, primary care centers in sections of the community with low physician-to-population ratios?[42]

In the present system of hospital finance and control there is no incentive even to ask such questions, because each hospital (or hospital chain) is an independent unit planning its own survival and expansion based on reimbursement for services it can market. The decisions made independently in one sector nevertheless have consequences for other sectors of the health system. For example, the provision of highly specialized, advanced technology in community hospitals encourages the further production of physician specialists by providing a marketplace for them to sell their services. The aggregate impact of that system is to focus as large an absolute and relative amount of resources as possible on the hospital sector; at the same time, hospital competition and pressures from physicians and medical technology companies encourage the even more narrowed

[41] Victor R. Fuchs, *The Growing Demand for Medical Care*, NEW ENGLAND JOURNAL OF MEDICINE 279(4):190-195, July 25, 1968.

[42] For an interesting attempt to examine the economic efficiency of alternative investments in a community's hospital-related resources, see Howard P. Greenwald, John M. Woodward, and David H. Berg, *Transportation or CT Scanners: A Theory and Method of Health Resources Allocation*, HEALTH SERVICES RESEARCH 14(3):207-219, Fall 1979.

focus of national health resources on ever higher levels of technological sophistication. In this way, resources are increasingly allocated to hospital capital independently of any societal assessment of the most appropriate use of those resources.

In all this expansion and acquisition of new technology, private hospitals had a distinct advantage over public hospitals, especially those 700-800 public hospitals that serve primarily low-income people. Although the Hill-Burton program attempted to put public and private not-for-profit hospitals on the same footing in their access to capital, private health insurance reinforced the differences between hospitals for the poor and those for the more affluent population. The growth of private hospitalization insurance assured voluntary hospitals of a secure and growing source of revenues. Public hospitals, which were effectively barred from serving privately insured patients, had much more limited revenues and resources. The situation was only exacerbated by the introduction of Medicare and Medicaid, which—despite the intentions of Congress—did not provide public hospitals with the same sources of patient care revenues as private hospitals.[43]

As a result, between 1950 and 1975 the supply of voluntary hospital beds increased 99%, while beds in public hospitals increased only 65%. In this period, however, public hospital assets per bed increased sharply, relying primarily on Hill-Burton funds and local tax resources. In 1950, public hospitals averaged only 65% of the assets per bed in voluntary hospitals, but by 1975 public hospitals had reached 87% of the asset level of voluntary hospitals.[44] Nevertheless, it is likely that a disproportionate share of this increase in public hospital assets was among those public hospitals serving a cross-section of their communities and those (primarily district hospitals in some states) serving only insured patients. Although it has not been well documented, there is evidence that, in physical plant, accommodations, and available medical technology, public and private hospitals that serve the poor are still well behind the larger portion of the hospital sector that caters to the more affluent and better-insured population.[45]

Thus, the gilded age of hospitals encouraged rapid expansion of the total supply of hospital beds and the acquisition of advanced medical technology. But it also created serious

[43] For an expanded discussion of the problems of public hospitals in a predominantly private hospital system, see E. Richard Brown, PUBLIC MEDICINE IN CRISIS: PUBLIC HOSPITALS IN CALIFORNIA, California Policy Seminar Monograph No. 11 (Berkeley: Institute of Governmental Studies, University of California, 1981).

[44] HOSPITAL STATISTICS, 1977 Edition, 5.

[45] See, for example, John C.M. Brust *et al.*, *Failure of CT Sharing in a Large Municipal Hospital*, NEW ENGLAND JOURNAL OF MEDICINE, 304(23):1388-1391, June 4, 1981.

fiscal problems for many hospitals, contributed substantially to the rapid increase in total health care costs, and encouraged the expansion of resources for expensive, high-technology care. Some of those resources might have gone for less expensive and more needed ambulatory care facilities and resources, particularly in the many urban and rural areas with designated shortages of physicians and other health services. Furthermore, some unknown portion of the large investment in construction, modernization, and replacement of existing facilities involved moving hospitals from inner-city poverty areas out to the suburbs or building additional new facilities in the suburbs and then drawing off resources from the aging inner-city hospitals. These developments were facilitated by the market system in which each hospital is a relatively autonomous producer and marketer of services, planning what market it will serve and the amount of expansion needed and desirable to maintain or expand its market share. In sum, the capital resources that have flowed into the hospital sector over the past three decades have reinforced the two-class hospital system and the trend toward highly specialized medicine, even as those resources have created benefits for that small portion of the population that truly needs such advanced technology.

Cost Containment and the Rationing of Hospital Care

The Glitter Fades. Until private insurance and government third-party payment programs became widespread, price rationing worked effectively, if inequitably, to constrain the total cost of hospital care. Individuals and families received a quantity of care commensurate, not with need, but with their incomes. Only the very affluent could afford to obtain the most complete and sophisticated care available. The growth of private insurance and its increasingly comprehensive coverage of hospital care eventually enabled the upper working class and middle class to get essentially the same care as the wealthier stratum of society. The introduction of Medicare and Medicaid provided similar opportunities for the growing proportion of the population who are elderly and for many of the nation's very poor. The socializing of payments for hospital care and the subsidizing of care for persons priced out of the insurance and hospital markets now cover more than 90% of total expenditures for hospital care. These plans and programs have enabled ever larger numbers of groups to participate effectively in the medical marketplace and, increasingly, to ignore the actual costs of the hospital services they receive. But these developments only explain the increasing economic demand for hospital care. The equally important issue, as far as increasing medical costs and the rationing of care are concerned, is how the hospital sector responded to that increased economic demand.

As we have seen, the market autonomy of hospitals, physicians, and the medical technology industries gives them a great deal of freedom to determine how to use those increasing economic resources injected into "their" system. A substantial portion of those resources have been used to expand the supply of hospital services, including the acquisition and use of expensive medical technology. This uncoordinated investment in new facilities and technologies and their unrestrained use produced a staggering increase in per capita costs of health care, with the costs per unit of service increasing considerably faster than the overall Consumer Price Index and with "intensity"—i.e., the kinds of goods and services and their utilization—contributing substantially to increasing costs. As the costs of health insurance rose to pay for the rising costs of hospital and other medical care, employers and individual subscribers mobilized pressure for some restraint imposed by Blue Cross and other insurers. But just as the entry of public subsidies into the hospital market contributed to even larger increases in hospital costs, the rapid drain of hospital costs on the Federal and state treasuries fostered the most serious cost-containment efforts.

Federal spending for hospital care jumped from $2.4 billion in 1965 (the year Medicare and Medicaid began) to $9.4 billion in 1970. By 1979, the Federal government was spending nearly $35 billion on hospital care, including nearly $22 billion through Medicare, $4 billion in Federal shares of the Medicaid program, and more than $7 billion on Veterans Administration and Department of Defense hospitals. State and local spending for hospital care grew less rapidly—from $3 billion in 1965 (80% of which went for their own public hospitals) to $12.8 billion in 1979 (with 53% going for care in public hospitals)—but it was an even greater burden on the more limited taxable resources at these levels of government.[46] With hospital costs the largest item in government spending for health care—in 1970, 65% of all pubic health care expenditures—and with the hospital industry showing no inclination to slow their rising costs, a political consensus quickly emerged that the government should step in to control the industry in which it had developed such a large stake.

Cost-Containment Strategies. Four approaches to rationing hospital care have emerged from these conditions.

Shifting the financial burden to patients. One method of reducing both the use of health services by beneficiaries and the consequent costs to the insurer is to make the patient responsible for some or all of the costs of care. This strategy forces the patient to weigh his or her need for care against its impact on available financial resources. Obviously, the lower a person's

[46] Gibson, *National Health Expenditures*, 1979.

income, the more a given charge for medical care will discourage seeking care. This approach has been applied to those sectors of the population for which government bears financial responsibility—namely, the elderly and the poor.

The Medicare program has from the beginning included deductibles, copayments, and exclusions adapted from private insurance as rationing devices. Even in 1977, with nearly all elderly Americans enrolled in the Medicare program, out-of-pocket payments by the elderly amounted to $3 out of every $10 of their total health costs and represented 11% of their per capita income.[47] Most of the elderly poor have been able to compensate for these cost-sharing restraints by being eligible for Medicaid as well. However, with congressional permission granted in Social Security amendments passed in 1967, 1971, and 1972, the states have gained increasing authority to impose copayments on Medicaid recipients, particularly for outpatient care. States are also free to reduce maximum income eligibility levels for some optional groups under Medicaid and to reduce or eliminate optional services; most states have done so.

In contrast to well-insured patients and more affluent population groups, low-income patients must decide whether their need for care is sufficiently great to warrant spending additional money or enduring greater difficulty or expenditure of time to obtain care. Accumulated health services research evidence suggests that very urgent care is less likely than preventive care or more moderately needed care to be reduced by modest cost-sharing. For example, the California experience with copayments for ambulatory care suggests that such policies may even result in higher overall costs to the Medicaid program, because more expensive services that require no copayment (such as hospitalization) tend to get substituted for less expensive ambulatory care that was deferred due to cost sharing.[48]

Restricting eligibility for Medicaid dramatically reduces the access of affected groups to private physician and hospital care. Poor persons who have no insurance coverage average substantially fewer physician vists each year than the poor with Medicaid.[49] The poor who have no third-party coverage have a difficult time finding a hospital that will treat or admit them, sometimes even for medically urgent conditions.[50] Private hos-

---

[47] Robert M. Gibson and Charles R. Fisher, *Age Differences in Health Care Spending, Fiscal Year 1977*, Social Security Bulletin 42(1):3-16, January 1979.

[48] Milton I. Roemer *et al.*, *Copayments for Ambulatory Care: Penny-Wise and Pound-Foolish*, Medical Care 13(6):457-466, June 1975.

[49] Davis *et al.*, *Access to Health Care for the Poor.*

[50] See, for example, *No Money, No Doctor: Deaths Show Poor's Need for Care*, National Health Law Program Health Advocate, No. 122:1, 4, July 1981.

pitals' requirements for pre-admission deposits, refusals by physicians to treat or admit indigent patients who are not eligible for Medicaid, and the virtual elimination of charity care from most private hospitals create enormous barriers for many poor persons in search of hospital care. Many Hill-Burton hospitals have been taken to court in order to get them to fulfill even their minimal obligation to provide emergency and inpatient care to indigent persons.[51]

These problems will become even more serious as profit-making hospital chains gain control of more hospitals and hospital beds. Although they have been studied very little, there is good reason to believe that for-profit hospitals skim the most lucrative patients and types of care from the hospital market and leave the rest to not-for-profit and public hospitals. The growing proportion of community hospitals that are for-profit—one in every eight nationally and one in every three in California—suggests that in the future the poor will have even greater difficulty in obtaining private hospital care.

The inability or refusal of private-sector medicine to care for the poor, especially those without Medicaid, forces the poor to rely heavily on public hospitals and clinics—in those communities that still maintain them. Restricting Medicaid eligibility therefore tends to shift the financial burdens of caring for the poor from the Federal and state governments to local governments, the level of government traditionally responsible for the poor and still responsible for publicly provided medical care. However, the increasing costs of providing that care, together with growing fiscal problems, have encouraged many local governments to increase charges for both inpatient and outpatient care and even in some cases to close their public hospitals. Local authorities have often justified their actions by declaring that the "truly" indigent are covered by Medicaid or Medicare, although this premise has not been supported by data on the Medicaid program (discussed earlier). In California, some county boards of supervisors have redefined indigency to coincide with the residency and income eligibility requirements of that state's Medicaid program, and they have then closed county public hospitals, arguing that all indigent people in the county were covered by Medicaid. California's once extensive network of 65 public hospitals operated by 49 counties in 1964 has shrunk to 37 hospitals in 29 counties in 1981.[52]

Thus, imposing cost sharing on patients and reducing eligibility and covered services in public programs shift the financial burdens of containing costs onto the poor themselves, without a doubt the group least able to bear the burden. Restrictions on Medicaid eligibility also shift financial burdens

[51] Freifeld, *Hill-Burton*.
[52] Brown, PUBLIC MEDICINE IN CRISIS.

from Federal and state governments to local taxpayers and governments, many of which have refused the responsibility by closing down their public hospitals and other health services. In the end, these rationing policies all redistribute the costs of medical care for the poor onto the poor themselves, reversing public policy and legislative trends over the past half-century. The emphasis on reducing inequities in health care has been replaced by an emphasis on controlling or reducing medical expenditures.

Controlling reimbursement to providers. Another approach to containing the costs of governmental third-party payment programs has been to hold down reimbursement rates to providers. This approach has generated two very different strategies.

One strategy, adopted by many states early in the Medicaid program's history, has been to maintain provider reimbursement rates at levels well below prevailing market rates. Since the original Medicare and Medicaid legislation requires that hospitals be reimbursed "reasonable costs" rather than set fees, the largest impact of restrictive reimbursement rates has been on the willingness of physicians to care for Medicaid patients.[53] Since Medicaid pays physicians 20-25% less than Medicare and Blue Shield, it is not surprising that the main complaint physicians make against Medicaid is what they regard as low reimbursement rates; nor is it surprising that most doctors treat few, if any, Medicaid patients. Just 6% of all physicians care for one-third of all Medicaid patients, and 20% of all physicians treat no Medicaid patients at all.[54] It is clear that simply keeping reimbursement rates below prevailing fees further reduces the ability of the poor to obtain care from private-practice physicians and increases their use of institutional sources of care—hospital outpatient departments, emergency rooms, and especially public hospital and clinics.

A second strategy to control reimbursement rates has been directed at hospitals and is generally more equitable, because it does not single out the poor with rationing policies that restrict their access. The pressures for cost containment have generated new reimbursement review procedures, guidelines, and controls on hospitals. Some states have established rate-setting agencies that group hospitals—for example, by size, urban/

---

[53] Stephen F. Loebs, *Medicaid—A Survey of Indicators and Issues*, in A.D. Spiegel, ed., THE MEDICAID EXPERIENCE (Germantown, Md.: Aspen Systems Corporation, 1979), 10-16.

[54] Ira L. Burney *et al.*, *Medicare and Medicaid Physician Payment Incentives*, HEALTH CARE FINANCING REVIEW 1(1):62-78, Summer 1979; Mitchell and Cromwell, *Medicaid Mills: Fact or Fiction?*; and Dewey D. Garner, Winston C. Liao, and Thomas R. Sharpe, *Factors Affecting Physician Participation in a State Medicaid Program*, MEDICAL CARE 17(1):43-58, January 1979.

suburban/rural location, teaching/nonteaching status, and the case mix of the hospital's patients—for purposes of comparing their costs; the agencies then disallow some costs for hospitals whose costs are excessive compared with those of similar hospitals. The mandate for rate-setting agencies and Medicaid fiscal intermediaries to examine, review, and allow or disallow costs has modified the interpretation of "reasonable costs" as the basis for reimbursing hospitals. Hospital costs are now scrutinized more closely than they were 15 years ago, and decisions as to what is "reasonable" are no longer made solely by the hospital administration.[55]

In addition to these reviews of hospital costs, new structures of reimbursement force hospitals to share fiscal risks with third-party payers. Some states, like New York and Massachusetts, have experimented with prospective reimbursement; in these states the hospital must live within rates set by the rate-setting commission in advance of the year for which the rates are paid. Others, like New Jersey, have adopted case-management reimbursement approaches in which a hospital is paid the same amount for each patient, or case, within a Diagnosis Related Group (DRG), with each hospital's reimbursement prospectively set between its own cost for caring for that DRG and the average cost for that DRG among comparable hospitals. A hospital whose cost is below the standard will reap a "profit," while a hospital whose cost is above the standard obviously will be reimbursed less than that cost. Both types of prospective reimbursement are intended to encourage economic efficiency on the part of the hospital, but there are serious concerns that excess revenues will be reinvested in the hospital and added to the cost base for future rates. In addition, it is unknown whether, in general, hospitals deal with excess costs by becoming economically more efficient, reducing staffing ratios to undesirably low levels, or cutting back on patient services.

Prospective reimbursement, whether restricted to public third-party programs or extended to private insurers as well, forces each hospital to bear the financial risks of cost-overruns. Fixed percentage limits on cost increases each year, as were proposed by the Carter administration, would similarly force hospitals to bear any risks of costs in excess of the permitted percentage increase. The problem with percentage limits on cost increases is that it penalizes lower-cost hospitals and successively widens the gap between them and higher-cost institutions.

Although most hospitals are in sufficiently sound financial positions to bear the risks, those public and private hospitals

[55] Stephen M. Weiner, *"Reasonable Cost" Reimbursement for Inpatient Hospital Services Under Medicare and Medicaid: The Emergence of Public Control*, American Journal of Law and Medicine 3(1):1-47, Spring 1977.

that are economically jeopardized because they remain in inner-city areas and serve primarily the poor are jeopardized further by having to bear additional financial risks. New Jersey's rate-setting system, however, reimburses hospitals that provide uncompensated care to indigent patients; it covers this cost by adding a surcharge to all third-party payers (including Medicaid), in proportion to their respective shares of the hospital insurance market.

If these reimbursement controls on hospitals are applied to private as well as public third-party payers, they do not themselves create any inequities. Unlike policies that maintain below-market reimbursement rates for practitioners, hospital reimbursement controls that affect all patients do not encourage hospitals to restrict access of Medicaid beneficiaries. Some programs, as in New Jersey, may even benefit the uninsured poor by removing the disincentive hospitals now face to providing charity care. Furthermore, by forcing hospitals to live within planned budgets, reimbursement controls may restrain the trend of hospitals' consuming an ever growing share of national health care resources.

Direct control of supply and provider behavior. A third approach to constraining hospital costs has been to impose controls directly on facilities and equipment and to limit the range of specific medical interventions. In this approach, agencies outside the medical institutions directly intervene to set explicit limits for capital investment in hospitals and to explicitly limit specific provider decisions, thus limiting provider autonomy.

One of the important forms of this approach is comprehensive areawide planning for health services, initiated nationally in 1966 with Comprehensive Health Planning agencies and succeeded in 1974 by Health Systems Agencies (HSAs), empowered to plan, review and help regulate the supply of health facilities and services in their areas. The objective of areawide planning is to coordinate the development and distribution of all health services to provide a medically effective and economically efficient health delivery system for each area—and, in effect, one that allocates equitably the community's health resources. Each planning body (including representatives of the "4 Ps"—the public, providers, payers and politicians) is to formulate an affirmative plan for development and then judge the present supply and proposed changes against the plan. Although such planning would in theory override the more parochial interests and market autonomy of individual institutions, hospitals have retained a disproportionately large share of power and control.

Planning is of minimal value without the power to implement one's plans. Certificate-of-need (CON) laws are the chief instrument of such power available to regional and statewide health planning agencies. State CON laws provide for legal

sanctions against hospitals that proceed with capital investment projects above a prescribed minimum dollar value (generally $100,000) without first having applied for and received certification of their community's need for the facility, equipment, or service. The Federal law, Section 1122 of the 1972 Social Security amendments, denies reimbursements by the Medicare and Medicaid programs for costs associated with a capital investment unless the investment has been approved by regional and statewide health planning agencies. CON programs have slowed expansion of hospital facilities, but they have had less effect on capital investment in equipment, modernization, and replacement. In general, CON proceedings have spawned a new specialty among law firms adept at preparing CON applications and using legal tools and political leverage to circumvent the legislative intent. CON proceedings have also tended to favor larger, wealthier hospitals, at least in part because of their greater access to legal and political resources.[56]

Despite their limitations, both areawide health planning and CON laws attempt to allocate health resources to the entire community, rationing them on the basis of health needs and correcting maldistributions based on per capita income, ethnic group composition, and urban/suburban/rural status. HSAs are limited in their ability to fulfil this mandate by (1) their political weakness—only provider representatives are effectively tied to organized political power outside the health planning apparatus; (2) the monopoly of technical planning resources among provider representatives vs. public members; and (3) the separation of planning from the purse strings—the power to regulate is feeble compared to the power to distribute money.

A second form of direct rationing is applied to the Medicaid practices of health professionals; for example, the use of drug formularies, limitations on the number of days of hospitalization or physician and dentist visits a patient may make each year, pretreatment review, and prior authorization for nonemergency hospitalization and for extensions of hospitalization beyond a specified maximum. These rationing devices are superimposed over professional judgment and thus represent a direct control over professional autonomy, in addition to the limits they place on the patient's own judgment concerning when and how often he or she needs to seek medical care. Although these devices are increasingly widespread among state Medicaid programs, they are strongly resented as bureaucratic intrusions by health professionals;[57] and they are particularly inequitable because they reduce the allocation of care to

[56] See Brust *et al.*, *Failure of CT Sharing in a Large Municipal Hospital.*
[57] Michael W. Jones and Bette Hamburger, *A Survey of Physician Participation in and Dissatisfaction with the Medi-Cal Program*, WESTERN JOURNAL OF MEDICINE 124(1):75-83, January 1976; and Garner *et al.*, *Factors Affecting Physician Participation in a State Medicaid Program.*

an already deprived group, reversing past redistributive efforts based on need. To the extent that such reviews and rationing devices reduce unnecessary treatment and are associated with more appropriate treatment, they can be credited with improving the quality of care provided to the Medicaid population. However, if and to what extent such improvements result, and under what review conditions, are matters of great contention among health care analysts and practitioners.

Direct control of the allocation process has thus been implemented in two different forms—limits on capital investment in hospitals and on health professionals' decisions. Both sets of methods seek to limit the market autonomy of providers—the major cause of the rapid escalation of medical costs, according to the analysis developed in this paper. However, methods that control the supply of hospital care attempt to ration care to the entire community on the basis of need, while methods to control clinical practice have so far been applied to only one population subgroup distinguished by financial dependence on a government subsidy program.

Indirect control of provider behavior. A fourth approach to containing medical costs has been indirect control of hospital administrator and health professional behavior by restructuring the environment in which they make decisions.

Hospital utilization review committees have had an impact beyond the specific cases to which their negative or corrective judgments have been made. They have also created an environment in many hospitals in which physicians, having received negative peer judgments in the past, try to anticipate such judgments by altering their clinical practice. Utilization review has probably exerted a modest influence on physicians' hospital practices through peer pressure, but it has avoided any fundamental changes in the structure of hospital or medical decisionmaking.

Prospective reimbursement of hospitals attempts to contain the cost of hospital care to third-party payers by creating a new fiscal structure within which hospital administrators make decisions. Capitation payments, which require the provider to operate within fixed revenues per subscriber during each year, go even further in restricting the incentive environment. As the payment basis for health maintenance organizations (HMOs) and other prepaid health plans, capitation payments create a disincentive to provide unnecessary care—in contrast to the fee-for-service system, which provides a monetary incentive for excessive diagnostic and therapeutic procedures (or at best no disincentive against unnecessary care). Even more than prospective budgeting, capitation payment systems would thus restructure the context for administrative and clinical decisionmaking within the hospital.

The major cost-saving factor in HMOs is a hospitalization rate lower than in conventional fee-for-service settings. The Southern California Kaiser Foundation Health Plan, for example, reported a target ratio of 1.8 beds per 1000 members in 1971,[58] a ratio maintained with a strong administrative and medical emphasis on alternatives to hospitalization, but also by rationing membership to younger and healthier groups than would be obtained from a cross-section of the population. Nevertheless, HMOs have been strongly advocated as a solution to the high costs of the Medicaid program on the premise that Medicaid recipients would receive high quality care at a lower cost to the public treasury, a premise that has seen little support from research evidence.[59]

These indirect methods of cost containment create an environment of resource limitations, in contrast to more traditional cost-plus reimbursement models, within which administrators and physicians must then make decisions. Because these methods create structural pressures on providers to make difficult allocation choices, rather than having external agencies specify what services should be provided or ordered for each patient, David Mechanic terms these indirect methods "implicit rationing."[60] Mechanic and many health policy makers favor implicit rationing because, compared with direct rationing, it retains a high degree of professional autonomy in clinical decisionmaking and a high degree of market autonomy for hospital administrators. Thus, while implicit rationing restructures the context of decisionmaking and in the aggregate may reduce overall hospital expenditures, it does not restructure decisionmaking itself within the hospital or general health care sector—a factor that makes it politically appealing but reduces its ability to deliver the significant improvements in quality of care and economic efficiency its proponents have claimed. Administrators and physicians are free to gear their services to economic demand, which may diverge widely from health needs.

Rationing Policies of the Reagan Administration. With a specific ideological commitment to reducing government spending for health and social programs, increasing military

[58] John J. Boardman, Jr., *Utilization Data and the Planning Process*, in A.R. Somers, ed., THE KAISER-PERMANENTE MEDICAL CARE PROGRAM: A SYMPOSIUM (New York: Commonwealth Fund, 1971), 65-66.

[59] Bruce Spitz, *When a Solution is Not a Solution: Medicaid and Health Maintenance Organizations*, JOURNAL OF HEALTH POLITICS, POLICY AND LAW 3(4):497-518, Winter 1979; and Richard E. Johnson and Daniel J. Azevedo, *Comparing the Medical Utilization and Expenditures of Low Income Health Plan Enrollees with Medicaid Recipients and with Low Income Enrollees Having Medicaid Eligibility*, MEDICAL CARE 17(9):953-966, September 1979.

[60] David Mechanic, *Approaches to Controlling the Costs of Medical Care: Short-Range and Long-Range Alternatives*, NEW ENGLAND JOURNAL OF MEDICINE 298(5):249-254, February 2, 1978.

spending, and unfettering the forces of the marketplace in many economic sectors, the Reagan administration plans to continue some recent trends in hospital rationing policies and reverse others.

One set of changes will shift more of the costs of medical care onto the poor. Cutbacks in the Medicaid program will be accompanied by the lifting of some restrictions on the states concerning eligibility, services covered, and cost sharing. These cutbacks and program changes will provide added Federal support to state policies that reverse past efforts to achieve equity in access to health services. The increased latitude granted to the states is also likely to result in expanded state control of specific Medicaid treatment decisions of physicians, a policy that is sure to reduce further the number of private-practice physicians willing to care for Medicaid recipients. Also, at least some states can be expected to enter into contracts with high-volume, relatively low-cost hospitals in each area to be the exclusive source of hospital care for Medicaid recipients in their areas.[61] This plan would aid public hospitals by channeling more Medicaid dollars to them, but it would also further channel the poor into an essentially segregated hospital system. These changes would move the society back toward price rationing of hospitalization insurance and hospital care and would reinforce the two-class system of medical care.

Second, the administration is eliminating federally mandated and supported programs that directly control the supply and behavior of providers. The phasing out of HSAs, Professional Standards Review Organizations (PSROs), and related agencies reverses the trend toward increased areawide health planning and regulation that began with the Hospital Survey and Construction Act of 1946 (the Hill-Burton Act) and culminated in the National Health Planning and Resources Development Act of 1974 (P.L. 93-641), establishing Health Systems Agencies and their related state and Federal agencies. Although state certificate-of-need laws are not directly affected by these Federal actions, it is likely that their implementation will be even less effective than at present without the HSAs, their staffs, and their consumer-majority boards—which have been the backbone of the programs in most states. It will now be up to individual third-party payers—the Federal Health Care Financing Administration (which is responsible for the Medicare and Medicaid programs), the Medicare and Medicaid fiscal intermediaries, Blue Cross, and other insurers—to restrain hospital capital investment by refusing to reimburse institutions for expenditures judged unnecessary or inappropriate. However, without areawide planning there will be no local criteria

[61] Linda E. Demkovich, *Choosing Hospitals for the Poor—Should the States Have That Power?* NATIONAL JOURNAL 12(37):1516-1518, September 13, 1980.

against which to judge specific proposals for capital investment. These changes will substantially enhance the market autonomy of hospitals, enabling them to roll back the clock toward the "good old days" before regulatory restraints were imposed.

Finally, the Reagan administration, like others before it over the past 15 years, will probably encourage certain forms of implicit rationing. Although the administration is eliminating subsidies for HMO development, it supports other policies that will create market incentives for HMOs. The support of the administration, and particularly of Office of Management and Budget director David Stockman, for a national program to encourage competition among health plans will probably generate an increased market for HMOs and thereby encourage their development. Without Federal development grants and subsidies, it is likely that most future HMOs will be profit-making ventures, obtaining needed capital from insurance companies and other large financial corporations. With the minimal regulation promised by supporters in the administration and in the Congress, market competition will probably lead to widely varying ranges of services, quality of care, and costs among the different plans. Under the "pro-competition" proposal, the upper middle and upper classes would undoubtedly pay the extra non-tax-deductible costs of "Cadillac" health plans, the lower middle and upper working classes would make do with more basic family models, and the lower working class and the poor would have to be content with stripped-down subcompacts (and hope their less affluent markets do not degenerate into the scandalous marketing and service conditions that characterized California's experiment with prepaid health plans for Medicaid recipients under then-Governor Reagan).

Other forms of implicit rationing—prospective budgeting and rate setting—will probably not be encouraged by the Reagan administration, but the administration will probably not prevent the states from continuing their own rate review and rate-setting programs. Nevertheless, the elimination of the Federal role in local and state health planning and the implementation of programs to encourage competition among health plans will greatly increase hospitals' market autonomy and thereby decrease the small degree of public accountability imposed on them in recent years. Relying on market forces to make hospitals and other providers responsive to the public means a return to price rationing, mitigated only by public subsidy programs that have in the past reinforced the division of hospital and medical care into a two-class system. Unless these policies and legislative programs are altered by the development of new political forces, the 1980s will be a decade of increased autonomy for hospitals and other medical industries, continued rapid growth of the for-profit hospital and HMO sector, and decreased equity for those priced out of the ever

expensive hospital and medical marketplace, reversing important social and political trends of the past half-century.

The Allocation of Resources for Medical Intensive Care

Albert G. Mulley, M.D., M.P.P.*

The high and rapidly rising costs of health care have brought to center stage some difficult economic questions. Have we reached or are we approaching an upper limit on what we are willing to spend for health? If so, how are newly constrained resources to be allocated efficiently? Can equity as well as efficiency be achieved?

Society's choices for health care expenditures are often posed in general terms. For example, should we be spending more for preventive programs and less on curative medicine? The questions become more difficult when the choices affect identifiable groups: Should we be spending more on hypertension control and less on chronic dialysis? These allocation decisions are most difficult when identified individuals are involved, when a clinical judgment about resource allocation can affect the likelihood of serious morbidity or death for specified patients: Who should be admitted to the last available bed in the intensive care unit—Mr. Smith or Mrs. Jones?

The purpose of this paper is to provide a case study of the process of allocating health care resources, focusing on medical intensive care. The study begins with a description of the development of intensive care in its many forms and traces the rapid diffusion of intensive care units (ICU) throughout American acute care hospitals. The current distribution of intensive care resources is examined. Evidence for and against the effectiveness of intensive care in its different forms and for different patients is briefly reviewed. The high and rapidly rising costs of intensive care and evidence of new constraints on those resources bring the study to the issue of rationing intensive care

*Medical Practices Evaluation Unit, Massachusetts General Hospital, Department of Medicine, Harvard Medical School, Boston, Massachusetts.

October 1981

resources at both the institutional and clinical levels.[1] In order to understand the form that rationing takes in the ICU, one must understand the kinds of patients admitted, their indications for admission, and the specific interventions that constitute "intensive care." After the demography of patient populations in two intensive care research units is reviewed, decision analysis constructs are used to characterize prototypical intensive care patients and to more clearly illustrate the clinical decisions that reflect real or perceived resource constraints. Decisionmaking roles at the administrative and clinical levels are examined to identify responsibility for rationing decisions. Evidence on current rationing of intensive care resouces is cited from the literature, and the basis for rationing at the clinical level is examined. Finally, the ethical issues of intensive care rationing and attendant policy implications are considered.

History, Availability, and Costs

Development of Intensive Care. An intensive care unit is an area of the hospital, usually with at least four beds, that is set aside for care of the most seriously ill. It is distinguished from the conventional ward by an array of electronic monitoring devices, by life-support machinery such as mechanical ventilators, and, perhaps most importantly, by a high concentration of nursing and support personnel. The nurse-to-patient ratio varies from ICU to ICU and from time to time within units, but it is often 1:1 and is rarely less than 1:3.

Intensive care units were first established for postoperative patients during the mid-1940s. By the early 1950s, such recovery rooms were reported by 20% of all community hospitals with 100 or more beds. The diffusion of such units, first through the largest hospitals, then to successively smaller institutions, was completed during the next decade. By the early 1960s, nearly all hospitals with at least 100 beds had established recovery rooms [Russell, 1979].

The special care offered by these early intensive care units seemed appropriate for other critically ill patients who had nonsurgical diseases. Seriously ill medical patients were admitted to early mixed (i.e., medical and surgical) intensive care units. But the burgeoning of medical intensive care followed the development of coronary care. Prompted by the availability of continuous electrocardiographic monitoring, electrical defibrillation, and other promising technological advances in the care of patients with complications of myocardial infarction, coronary care units (CCU) were introduced in the early 1960s [Day, 1963; Brown *et al.*, 1963]. The rationale for these early

[1] For the purposes of this study, rationing is defined as limiting the consumption of medical care so that a patient does not receive all the care that he or his physician believe would be of some benefit to him.

CCUs was to provide a setting that insured immediate availability of equipment and trained personnel to respond to complications, particularly life-threatening arrhythmias. It was hoped that intensive monitoring and immediate care for those who developed complications could significantly reduce the 30-40% hospital mortality of patients with acute myocardial infarction. With experience, the CCU goal was extended to include the prevention of major complications by providing early treatment for minor arrhythmias and hemodynamic abnormalities [Lown *et al.*, 1967]. More recently, CCU therapy has been directed at altering the course of patients with acute ischemic heart disease and minimizing the extent of myocardial damage.

As intensive care units and coronary care units became commonplace, intensive care became a general response to critical illness. High mortality rates among medical as well as surgical patients requiring respiratory support prompted the introduction of respiratory intensive care units [Rogers *et al.*, 1972]. Similar beds were made available to patients with pulmonary edema, gastrointestinal hemorrhage, and other acute conditions [Callahan *et al.*, 1967; Lepore and Grace, 1969; Griner, 1972]. Highly specialized units designed to deal with the problems of burn victims and of high risk neonates were established and became widely distributed. By 1977, 99% of short-term general hospitals with 200 or more beds responding to the annual American Hospital Association survey had intensive care units. Of the nearly 3000 hospitals with at least 100 beds, 95% reported intensive care units [American Hospital Association, 1978]. It is apparent that today the intensive care unit is generally considered highly effective by the physician and is viewed by the patient as evidence of technologically advanced, competent care. The providers and consumers of care now expect such units in all but the smallest community hospitals.

Distribution. The geographical distribution of intensive care and coronary care unit beds by census region in 1976 appears in Table 1 [Russell, 1979]. The total number of beds is related to census region populations as well as the number of hospital beds in each region. West North Central and Pacific regions had a higher number of ICU and CCU beds per 10,000 persons, while Middle Atlantic and East North Central regions had relatively fewer beds designated for intensive care. However, the distribution is surprisingly uniform. More recent data indicate that the number of ICU/CCU beds has decreased (by as much as 18%) where they were relatively scarce [American Hospital Association, 1980].

Within regions, the diffusion of intensive care technology between 1961 and 1975 was not uniform [Russell, 1979]. The variable speed of adoption of intensive care by different kinds

TABLE 1: Geographical Distribution of Intensive Care and Coronary Care Beds

1976 U.S. Census Region	Number of Intensive Care and Coronary Care Beds		
	Total	Per 10,000 population	Per 100 hospital beds
New England	2,928	2.4	5.3
Middle Atlantic	8,213	2.2	4.5
South Atlantic	7,894	2.3	4.8
East North Central	8,805	2.2	4.3
East South Central	3,180	2.3	4.4
West North Central	4,731	2.8	4.5
West South Central	5,063	2.4	4.6
Mountain	2,253	2.3	5.2
Pacific	8,003	2.8	6.6
United States	51,070	2.4	4.8

Source: L.B. Russell, TECHNOLOGY IN HOSPITALS: MEDICAL ADVANCES AND THEIR DIFFUSION (Washington, D.C.: The Brookings Institution, 1979).

of hospitals provides an historical perspective on the question of availability of intensive care to different patient populations.[2] Larger hospitals were more likely to establish ICU/CCUs early, as were hospitals with medical school affiliations and/or residency programs. Private nonprofit hospitals adopted ICUs earlier than private, for-profit hospitals or those run by state or local governments. There was little relationship between speed of adoption and crude measures of case-mix such as the rate of heart disease deaths or motor vehicle deaths adjusted for the number of hospital beds. Hospitals serving a high proportion of nonwhite patients were likely to establish ICUs later than those serving predominantly white populations.

In 1975, the proportion of hospital beds dedicated to intensive care varied widely among hospitals, ranging from 0-40% with a mean of 5%. Not surprisingly, larger hospitals had more ICU beds, but the increase was not proportional to hospital size. Hospitals with 100-199 beds dedicated the greatest proportion of beds to intensive care. Hospitals with medical school affiliations and those with major investments in postgraduate educa-

[2] Throughout this paper, the term "availability" is used to refer to the proportion of times that a bed is accessible to a patient when it is considered of potential benefit and is requested by a physician.

tion had relatively more ICU beds. Those supported by state and local governments and those serving a high proportion of nonwhite patients had relatively fewer ICU beds. There was no clear relationship between proportion of ICU beds and the deaths per hospital bed due to heart disease or to motor vehicle accidents.

In summary, intensive care appears to be quite evenly distributed geographically. Patients served by larger, nonprofit hospitals with medical school affiliations and residency programs were the first to have access and still appear to have greater ICU bed availability. Patients cared for in public hospitals serving many nonwhite patients were the last to have access; it appears that ICU bed availability in these hospitals continues to be relatively low. There is little evident influence of hospital case-mix on either readiness to adopt intensive care or the proportion of beds dedicated to intensive care.

Effectiveness. Intensive care units have gained wide acceptance despite unproven effectiveness. The conviction that ICUs decrease morbidity and mortality is supported only by uncontrolled or poorly controlled studies. Many studies have attempted to document CCU effectiveness [Killip and Kimball, 1967; Oliver *et al.*, 1967; Klaus *et al.*, 1970; Hofvendahl, 1971; Christiansen *et al.*, 1971]. Most have demonstrated lower myocardial infarction case-fatality rates for patients cared for in a CCU than for those cared for in a general medical ward.[3] However, these studies purporting to demonstrate CCU effectiveness have compared unit mortality rates with historical controls or with patients nonrandomly assigned to conventional care. Obviously, serious questions can be raised about patient selection, diagnostic criteria, and many potential confounding variables. Other studies have failed to demonstrate an effect of CCU care on the myocardial infarction case-fatality rate. The real impact of coronary care or even hospitalization for a subset of patients with myocardial infarction has been questioned by randomized studies in Great Britain that have found no difference in outcome for selected patients who are treated at home and those who are admitted to the hospital [Mather *et al.*, 1976; Hill *et al.*, 1978].

Attempts to evaluate the effectiveness of intensive care for noncoronary patients have similarly relied on comparisons with historical controls and patients nonrandomly assigned to conventional care. Reductions in respiratory failure mortality rates have been shown after the introduction of respiratory intensive care units [Rogers *et al.*, 1972]. Studies of stroke patients suggest that outcome is not altered by intensive care

[3] A review of comparative studies, all of which used concurrent but not randomized controls, found a mortality rate of 21% for 2492 CCU patients and a mortality rate of 32% for 1648 general ward patients with myocardial infarction.

[Pitner and Mance, 1973]. In a study of patients with pulmonary adema, Griner found significantly higher rates of diagnostic and therapeutic intervention, but no difference in outcome after the introduction of an intensive care unit [Griner, 1972]. In the same hospital, intensive care similarly did not affect the mortality rate for patients admitted after intentional drug overdose [Piper and Griner, 1974].

For some observers, then, the health benefits of intensive care remain uncertain. Nevertheless, physicians' belief in the value of intensive care is not an article of blind faith. While the effectiveness of care is unproven, the efficacy of many of the individual interventions that can be practically applied in an ICU setting is without doubt.[4] Respirators do maintain ventilation, pacemakers do maintain heart rhythms, pharmacologic agents do control blood pressure, transfusion does sustain oxygen-carrying capacity, and dialysis does eliminate potential toxins.[5] Furthermore, the common denominator of intensive care, in different settings and for different patients, is closer attention to patients' condition and therapy. While the potential for patient discomfort and iatrogenic disease in the ICU is well-recognized [Hackett *et al.*, 1968; Steel *et al.*, 1981], it is difficult to deny the potential benefit of more attention rather than less for many patients with critical illnesses or at risk for them. It is, therefore, not surprising that those who make intensive care allocation decisions for individual patients generally view the benefits as unmeasured rather than uncertain.

Costs. During the past 25 years, the percentage of the gross national product (GNP) consumed by the health care enterprise has doubled. The health care share is rapidly approaching 10%. These rapidly rising costs have focused attention on expensive high-technology care such as that provided in the ICU [Griner, 1973; Mass. Physician, 1975]. While costs for medical services were rising at a rate of 9.5% during the mid-1970s, hospital costs, which account for 40% of the health care bill, have been going up at an average annual rate of more than 17%. In Massachusetts, total costs of non-Federal acute hospitals grew at a rate of 15% per year from 1973-1976. During this period, the costs of intensive care units were rising more rapidly than any other category of hospital costs, at an annual rate of 28% [Shepard and Ghanotakis, 1979]. This increase in cost rates far

[4] The distinction between efficacy and effectiveness is noteworthy. An intervention is *efficacious* if it *can* provide health benefits when applied under ideal circumstances. It is *effective* if it *does* provide benefits when used under normal circumstances.

[5] Efficacy has not been demonstrated by randomized controlled trials; but because survival is unprecendented without respiration, heart rhythm, blood pressure, blood volume, or renal function (natural or artificial)—and because survival does occur with these interventions—such trials are not necessary to demonstrate efficacy for some patients.

outstripped the annual increase (less than 5%) in the number of intensive care beds. A review of individual hospital cost data discloses that it is not uncommon for intensive care units to account for as much as 20% of total hospital costs. The number of beds dedicated to intensive care rarely exceeds 10%.

A number of economists have estimated the proportion of total hospital expenditures that can be attributed to intensive care. [Russell, 1979; Wagner, 1980a.] Such estimates, however, must rely on hospital charge information[6] and uncertain estimates of the ratio of intensive care charges to conventional care charges. Hospital charges for intensive care vary widely, even within institutions. The ratio of per diem intensive care charges to conventional care charges reported in the literature range from 2.0 to 5.0. Presumably this variability reflects differences in case-mix, in the types of interventions commonly made within the unit, and in staffing requirements as well as the degree of cross-subsidization within the hospital. The ratio of ancillary charges incurred by patients in ICUs to those incurred by patients receiving conventional care is even more uncertain and has been estimated to be in the 1.0-3.5 range. Using these estimates, together with empirically based estimates of occupancy rates for ICU beds and conventional beds, Wagner has estimated that 15.2% of hospital expenses are for adult ICU patients. The ICU share climbs to 16.8% when neonatal intensive care beds are included at the same relative charge ratios and occupancy rates. These percentages amount to $8.9 billion and $9.8 billion respectively, of the $58.348 billion dollars spent in all acute non-Federal hospitals in 1978 [Wagner, 1980a]. During the late 1970s both the proportion of ICU beds and the ratio of ICU to conventional care costs have been increasing. The result has been even higher proportions of hospital costs being expended in the ICU.

It must be kept in mind that estimates of total intensive care costs may not provide a clear measure of the economic impact of the ICU. Diagnostic and therapeutic modalities, and their costs, have expanded regardless of the setting in which they are used. For example, the costs of caring for patients with acute coronary disease were already rising rapidly before the introduction of coronary care units [Scitovsky, 1967; Martin *et al.*, 1974].

Theoretically, bringing patients with special needs and personnel with special skills together in special units should have introduced efficiencies and lowered costs. Griner has shown, however, that a higher level of diagnostic and therapeutic intervention is more likely for patients admitted to an ICU than for

[6] There is evidence that due to interdepartmental cross-subsidization within hospitals, intensive care charges underestimate actual costs [Harris, 1979]. However, it is likely that the degree and direction of cross-subsidization vary from one hospital to another.

patients with similar outcomes treated on a general medical ward [Griner, 1972]. It follows that one explanation for the high costs of intensive care is the ease with which interventions, including those with marginal expected benefits, can be made when the technology and personnel are readily available.[7]

ICU Patients

The dramatic increase in hospital resources dedicated to intensive care forces us to consider economic questions. How are decisions made to allocate resources to intensive care within a hospital or planning region? Are available resources insufficient, sufficient, or excessive to meet the needs of those likely to benefit? How were intensive care resources rationed among patients likely to benefit in the past when fewer resources were available? How often must such rationing decisions be made today despite increased availability? We can better consider the evidence that bears on these questions after first describing the currently accepted indications for admission to an intensive care unit, which can be inferred from descriptions of ICU popultions.

Diagnoses and Demographic Characteristics. In order to better characterize patients receiving intensive care in this age of apparent wide availability of ICU beds, two descriptive studies recently were initiated independently—one at the Massachusetts General Hospital (MGH) [Thibault *et al.*, 1980], and the other at the George Washington University Hospital Center (GW) [Knaus *et al.*, *Range of Intensive Care Services*, 1981]. The studies are complementary. The ICU at MGH serves patients with critical medical illnesses, including coronary artery disease. A majority of patients admitted to the ICU at GW have surgically treated diseases; patients with acute coronary artery disease are excluded. Together these two studies encompass the range of adult patients receiving intensive care in medical center hospitals.

Massachusetts General Hospital. The facility studied is an 18-bed combined coronary and medical ICU. Six other areas within the hospital provide intensive care to adult patients. Patients previously followed by, or specifically referred to, staff physicians are generally admitted to either a 6-bed CCU or a 10-bed ICU elsewhere in the hospital. Surgical patients and occasional nonsurgical patients with respiratory insufficiency are cared for in one of the three ICUs staffed by the departments of anesthesia and surgery. A combined surgical and medical ICU for patients with neurologic diseases is staffed by the neurology and neurosurgical departments. The availability of these alter-

[7] The physical availability of equipment and personnel may be only part of the answer. ICU admission may represent a first step that reinforces a commitment to a course of aggressive intervention.

native intensive care sites within the hospital introduces some selection biases; medical patients referred from outside hospitals for tertiary care and those presenting with evident neurological disease are likely to be underrepresented in the study ICU. Nevertheless, the patient population of the study unit is otherwise believed to be representative of general medical ICUs serving urban areas. More than 85% of patients admitted lived within a 15-mile radius, and only 11% of admissions were in transfer from another hospital or referred from an outside emergency room.

During the first two years of the MGH study (1977-1979), 2305 patients were admitted to the ICU 2693 times. Of these admissions, 58% were male patients, and 42% female. The mean age was 60 years (median, 62). The mean age among female patients was four years older than that among male patients. Age distributions for each sex [Figure 1] indicate that men predominated between 30 and 70 years—strikingly so during the fourth, fifth, and sixth decades.

Approximately 85% of patients were able and willing to have an interview that allowed a detailed assessment of their preadmission physical capacity, functional status, and of their perceptions of their health. Most patients who were not interviewed were either intubated, comatose, or otherwise too sick.

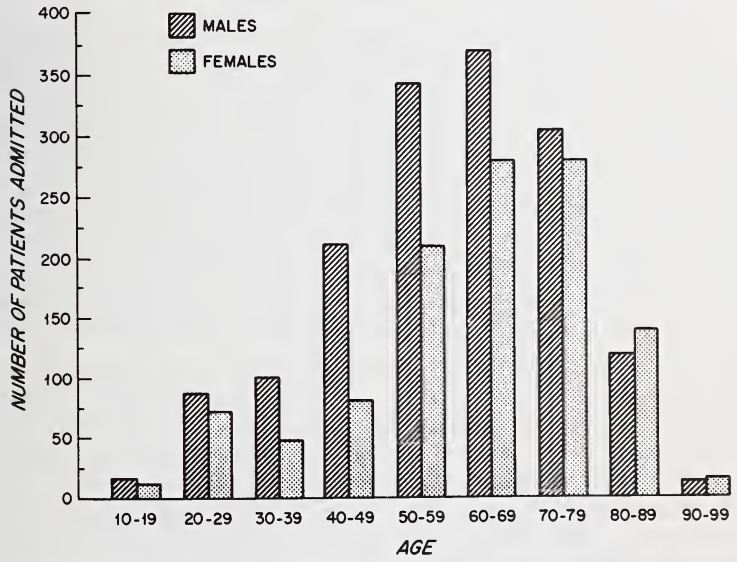


Figure 1. Age Distributions among Male and Female Patients with 2693 Admissions to the ICU/CCU

Source: G.E. Thibault, A.G. Mulley, G.O. Barnett, *et al.*, *Medical Intensive Care: Indications, Interventions, and Outcomes*, N Engl J Med 302:938-942 (1980). Reprinted with permission.

About half the patients who could not be interviewed died in the hospital. More than half the patients who were not working full-time attributed their employment status to poor health. The patients' perception of their health varied little with age; "considerable" and "severe" health-imposed limits on quality of life were reported by 25% and 16% of patients, respectively. Most patients were intensive consumers of medical care as well as consumers of intensive care. The mean number of outpatient visits to a physician during the year preceding admission was seven (median, three). Nearly half of all patients were also hospitalized at least once during that same period.

Suspected myocardial infarction on the basis of presentation with chest pain was the most common primary admission diagnosis, accounting for 39% of all admissions.[8] Respiratory diseases including pneumonia, chronic obstructive pulmonary disease, asthma, and respiratory failure accounted for 11% of all primary admission diagnoses. The third most common problem, leading to 6% of all admissions to the ICU/CCU, was intentional drug overdose.

Mean ages and sex ratios for each admission diagnosis indicate that the predominance of men among patients in the fourth through seventh decades can be explained by a higher incidence of cardiovascular and respiratory disease among men in that age group. Drug overdose was the most common diagnosis among patients less than 50 years old.

George Washington University Hospital. The ICU studied at GW is a 16-bed mixed medical-surgical unit that provides the only intensive care for the 500-bed hospital with the exception of a separate 8-bed coronary care unit. The ICU population at GW is therefore likely to be representative of mixed units (excluding coronary care) found in any large hospital [Knaus *et al.*, *Range of Intensive Care Services*, 1981; Wagner, 1980b]. During the first eight months of the GW study, 613 patients were admitted to the ICU, 54% male and 46% female. The mean age was 53 years, and 26% of patients were over age 65. At GW, 55% came to the ICU from the operating room; remaining patients were most often admitted from the emergency room or transferred from the hospital wards because of cardiovascular, neurologic, or respiratory failure or insufficiency [Draper *et al.*, 1981]. Of the ICU patients, 24% were in failing health (unemployed because of illness, bedridden or severely limited by symptoms, or with one or more weekly physician visits); an additional 14% were not as restricted but had serious chronic health problems [Wagner, 1980b].

[8] A single "primary" admission diagnosis was assigned on each admission. Secondary diagnoses were also recorded. Suspected myocardial infarction was an indication (primary or secondary diagnosis) for 58% of all admissions.

Intensive care units serve widely varying populations; generalizations from these observations must be made with caution. It appears, however, that the elderly and those with chronic illnesses consume a substantial proportion of ICU resources. Coronary disease and prolonged monitoring and support of postoperative patients are the most common reasons for ICU admission.

Need. The diagnoses and demography of current ICU populations shed some light on the accepted indications for intensive care. However, the allocation of ICU resources is determined by individual clinical decisions to admit a patient to an ICU, to maintain the patient in that ICU, or to transfer the patient at some point in the hospital course from the ICU. We can better understand these allocation decisions by first considering prototypical clinical problems faced by ICU physicians. The components of the decisions can be more explictly illustrated using decision analysis constructs. See Figure 2 [Mulley, 1980].[9]

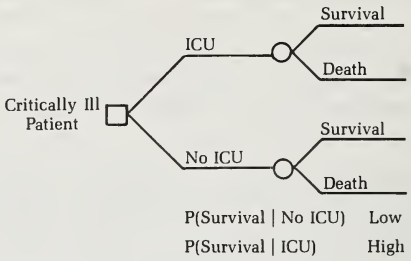
Case 1: The patient with a good prognosis with intensive care and a bad prognosis without intensive care. The first case (Figure 2a) involves the simplest decision—whether or not to initiate or continue intensive care for a patient with acute reversible disease for whom the probability of surviving without such intervention is low and the probability of surviving *with* such intervention is high. A common clinical example (accounting for 6% of all admissions to the ICU/CCU at MGH) is the patient with respiratory failure secondary to drug overdose. This decision tree and the class of decisions it represents may seem trivial. The message is that when there is clear conviction that ICU intervention is efficacious—that the probability of a good outcome given ICU intervention is clearly greater than the probability without such intervention—then the decision to intervene is straightforward. Rarely under these circumstances do physicians consider the increment in resource costs necessary to provide intensive care. Resources are not withheld from these patients except under unusual circumstances.

A number of simplifications should be noted in this model. First, intensive care is treated as if it were a single intervention when, in fact, it is a composite of interventions about which individual decisions can be made. For example, the overdose victim may be intubated and given ventilatory support even if other components of usual intensive care are unavailable. Second, the model is static; it doesn't adequately reflect the time dimension of ICU decisionmaking unless it is

[9] The "decision tree" simply illustrates the elements of a decision: the choices, represented by branches emanating from the square nodes; the contingent events that happen by chance, represented by the branches emanating from the round nodes; and the outcomes of concern to the decisionmaker, represented by the terminal branches.

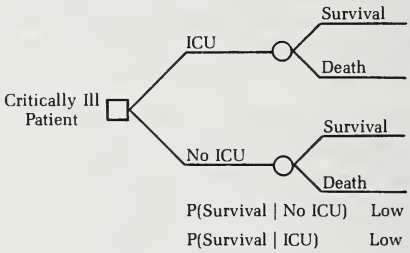
Case 1. Good Prognosis With ICU Intervention

Figure 2a



Case 2. Poor Prognosis Despite ICU Intervention

Figure 2b



Case 3. Good Prognosis With or Without ICU Intervention

Figure 2c

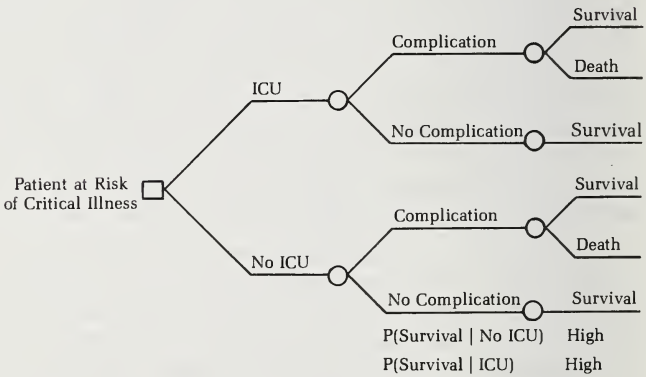


Figure 2: Prototypical ICU Admission and Transfer Decisions

Source: A.G. Mulley, *Intensive Care Decision Making*, presented at the annual meeting of the Society for Medical Decision Making, Washington, D.C., 1980.

viewed as one in a series of snapshots. The relative probability of survival given intervention may change from day to day or even moment to moment. The decision to continue an intervention is reassessed as the clinical course evolves and new prognostic information becomes available. The third simplification regards outcomes—which in reality are not dichotomous. Intermediate outcomes, including survival of varying duration and various degrees of morbidity, would be included in a more realistic model.

Case 2: The patient with a bad prognosis with and without intensive care. The second example has the same structure (Figure 2b). However, while the probability of survival is again low for patients without the benefit of ICU intervention, it is also low, perhaps equally so, for those for whom intervention is provided. Clinical examples include the patient with gastrointestinal bleeding and underlying hepatic failure or the patient with respiratory failure superimposed on chronic pulmonary disease. Again, this decision tree is too simple—particularly at the outcome end. Among possible outcomes for patients with chronic, irreversible disease, the best will include substantial morbidity; death may not be the least valued result for some patients in certain circumstances. Again, decisions may be reassessed as new information becomes available and the clinician becomes more certain about his probability assessments.

This decision tree represents the kind of ICU decision that has attracted the most attention from those concerned about the dollar and human costs of intensive care. A number of investigators, most notably Cullen and colleagues, have documented the high costs and high morbidity and mortality that can result from intensive care [Cullen *t al.*, 1976]. Focusing on the 226 most critically ill patients admitted to a primarily surgical intensive care unit, they found hospital charges averaging nearly $15,000; by 12 months after admission only 23% of the patients had been discharged alive and only 12% had fully recovered.[10] The authors concluded: " . . . cost control can be achieved in two ways: by not accepting patients for whom intensive care is inappropriate and by discontinuing intensive care in patients whose survival to a successful outcome is highly unlikely."

Similar observations have been made in different settings [Civetta, 1973; Morgan *et al.*, 1973; Turnbull *et al.*, 1979]. In each case, similar conclusions favoring "a selective admission and discharge policy . . . to concentrate available resources on those with a realistic chance of survival" [Turnbull *et al.*, 1979]

[10] These charges are from the early 1970s. Adjusted for hospital charge inflation, the average charge would be in excess of $75,000 in 1981 dollars.

were drawn. Committees have been formed and guidelines established to assist physicians in making decisions about continuing intervention in hopelessly ill patients [Pontoppidan *et al.*, 1976; Tagge *et al.*, 1974]. Such guidelines provide classification schemes that relate prognosis and effectiveness of available therapy to a schedule of intervention. While such schemes enable more explicit and simplified communication among care givers, difficult judgments are still required from physicians. For example, the decision to withhold any active treatment from a patient in the surgical ICU at Mount Sinai Hospital in New York requires the judgment "that his estimated chances of survival are of the order of magnitude of one in 100,000; that he has not responded to active, aggressive and presumably maximal therapy; and that the nature of his disease precludes a satisfactory quality of life even with survival." [Tagge *et al.*, 1974.][11]

Case 3: The patient with a good prognosis with and without intensive care. The third example is illustrated by a more complex decision tree (Figure 2c). The patient is admitted to or kept in the ICU not because he is critically ill but because he is at risk of becoming critically ill. The purpose of intensive care in many of these instances is not so much to prevent a serious complication—for the probability of a complication may not be substantially different with or without intensive care—but to allow a prompt response to any complication that may occur. It is presumed that the prompt response to a potentially fatal complication made possible by continuous monitoring in the ICU increases the probability of a favorable outcome.

This model represents the rationale for the introduction of coronary care units. It also represents the most common indication for admission to medical or surgical intensive care units. Among all patients admitted to the MGH medical ICU/CCU, 77% did not require any major diagnostic or therapeutic interventions at the time of admission and were admitted only for monitoring [Thibault *et al.*, 1980]. Among patients admitted to the GW medical-surgical ICU, 49% were admitted for monitoring rather than active treatment [Knaus, *et al., Range of Intensive Care Services*, 1981].[12] In general, the probability of com-

[11] When formal criteria for withholding therapy are this stringent, it could be argued that the decision is based solely on a judgment about the patient's welfare; that is, when morbidity is so great and irreversible, death may be preferable. Such a decision would not constitute rationing in the sense of denying resources with presumed benefits. However, the Mount Sinai Hospital committee did note that "such decisions allow for more attention to be appropriately redeployed to the other patients with greater benefit."

[12] While this may reflect a difference in disease severity in the two units, it also reflects the difference between medical and surgical ICU admissions. Many patients admitted after surgery remain endotracheally intubated and would therefore receive active treatment in the form of mechanical ventilation.

plications requiring major intervention is low among patients admitted for intensive monitoring; 10% of the 77% at MGH and 14% of the 49% at GW developed the need for a major intervention during their ICU stays. As a result, the probability of a good outcome is high with and without the ICU. Specific clinical examples of this prototype include patients admitted to medical ICUs and CCUs with suspected myocardial infarction after presentation with uncomplicated chest pain [Mulley *et al.*, *The Course of Suspected Myocardial Infarction*, 1980] and patients routinely admitted to surgical ICUs following a neurosurgical procedure [Knaus *et al.*, *Neurosurgical Admissions*, 1981].

Compared with use of intensive care resources for patients represented by cases 1 and 2, the expenditures of intensive care resources for patients admitted or maintained in ICUs only for monitoring can be considered a luxury aimed at reducing already low morbidity and mortality rates to even lower levels. Because the rates are low and perhaps because there is some distance, both temporal and causal, between the decisions and their consequences, rationing is most evident in the allocation of intensive care resources to these patients. The benefit being rationed is the presumed reduced probability of a bad outcome if a feared complication occurs.

Taken together these models represent all ICU decisions, but the representation is too neat. It must be kept in mind that the probability assessments incorporated in each of these decision trees are often made with a great deal of uncertainty. It may not even be clear which prototype is appropriate for a specific patient at a particular point in time.

Rationing

Decisionmaking. The burgeoning growth of ICU beds and intensive care costs during the past decade would suggest that institutional decisionmakers in hospitals and planning agencies recognize few resource constraints in determining the availability of intensive care. During this period of growth, many formulas have been devised for calculating both institutional bed need and areawide bed need [Cooper and Corcoran, 1974; Arthur D. Little, Inc., 1979]. Such formulas generally require an estimate of the number of patients requiring intensive care, the average ICU length of stay,[13] and either an optimal ICU occupancy rate or an acceptable rate, arbitrarily chosen, of ICU bed availability.

Because of the trade-off between occupancy and availability,[14] ICU bed needs have most often been calculated using

[13] Note that these estimates are usually empirically derived and therefore assume current utilization of ICU resources to be "appropriate."
[14] The trade-off is steeper for small units; that is, a small unit will have a lower availability rate than a larger unit with the same occupancy rate.

availability rates between 80 and 90%. Such a policy tacitly accepts the constraints on intensive care resources and determines that 10-20% of potential ICU admissions either will be rejected or will displace a patient presumably thought to be benefitting in some way from his position in the ICU.

The trade-off between occupancy and availability underlies an internal conflict in the hospital between administrators and physicians. Hospital revenue increases with increased occupancy. Doctors care far less about the costs of maintaining a certain capacity and far more about the availability of an ICU bed for their patients when the need arises [Harris, 1975].

The control that physicians can exercise in determining capital expenditures and hospital capacity has been debated; it probably varies among institutions. It has been argued that since most administrators and trustees feel ill-prepared to determine medical needs, large capital expenditures are often made as a result of physician advocacy [Redisch, 1978].

There is no doubt that physicians are the key decision-makers on the demand side within the hospital. Relatively subjective decisions about patients' medical needs determine demand for intensive care as well as other services and are institutionalized in the need formulas of hospitals and planning agencies.[15]

These institutional decisions, variably influenced by physicians' perceptions of medical need, determine short-term availability of intensive care resouces. Availability, in turn, determines the degree of rationing among patients. The level of expected benefits for an individual patient necessary to justify intensive care admission or continuing intensive care will be greater in hospitals or in areas where availability is low.[16] In settings where availability is low, the following results might be expected: higher rates of major interventions, case-mixes, and mortality rates reflecting greater disease severity; and higher rates of complications following transfer from intensive care, prompting readmission.

The change in composition of CCU patient populations during the past 15 years provides an historical example. Nearly all patients admitted to early CCU beds were admitted with a definite diagnosis of acute myocardial infarction. As CCU beds became more available, patients with less severe forms of coronary disease were more likely to be admitted. Despite increased

[15] Recall that Russell's study would support the view that physician behavior and influence may be a more important determinant of availability than institutional case-mix.

[16] This assumes that if any level of benefit is expected, and ICU beds are available, they will be used. Evidence for this ICU version of "Roemer's law" (i.e., that there is no practical limit to the number of beds that, if available, can be filled with patients) is presented later.

sensitivity of diagnostic tests, the incidence of myocardial infarction among CCU patients admitted with suspected myocardial infarction has declined steadily in the past 20 years and is now less than 30% in some units [Brown *et al.*, 1963; Killip and Kimball, 1967; Lown *et al.*, 1967; Bloom and Peterson, 1973; Virginia RMP, 1975; Eisenberg *et al.*, 1979; Thibault *et al.*, 1980].[17]

A comparison study done by the ICU Research Unit at GW provides another example of the manner in which availability can affect the stringency of rationing decisions [Draper *et al.*, 1981]. Recall that the ICU at GW is a 16-bed mixed medical-surgical unit serving a 500-bed hospital that is a major referral center. The ICU population at GW was compared with that of a 20-bed unit serving a community hospital with a total of 300 beds. Both hospitals had separate CCUs. During the eight-month period, 23 admissions were denied to the ICU at GW primarily because of a shortage of ICU nurses. No patient whose admission had been requested by his attending physician was denied ICU treatment at the community hospital. The units had the same technical and personnel resources with similar nurse-to-patient ratios varying from 1:1 to 1:3. A greater proportion of ICU patients at the community hospital were over age 65 (44% vs. 26%), and more patients tended to be in poor or failing health prior to admission. Level of intervention measured by the therapeutic index scoring system [Cullen *et al.*, 1974], severity of illness measured by an acute physiology score [Knaus *et al.*, *APACHE*, 1981], and mortality rates were significantly greater in the ICU at GW.

Recently, shortages of intensive care nurses have caused sudden decreases in ICU bed availability.[18] Such a reduction at Massachusetts General Hospital has resulted in an apparent increase in the rate of major interventions among patients admitted to the remaining open beds. Patients previously admitted to ICU beds for intensive monitoring have been admitted to conventional care beds. In some settings higher rates of serious complications occurring after transfer from the ICU and resulting in death have been documented during periods of reduced ICU bed availability due to nursing shortages [Franklin *et al.*, 1981]. On the other hand, in at least one coronary care setting,

[17] Practices in the first CCUs also tell us something about the importance of affinity for a known patient and of commitment to a particular course of action. Many with infarction remained in the CCU throughout their hospitalization, while, presumably, other patients were denied access.

[18] The shortage of nurses has become particularly critical in intensive care units. ICU staffs face special problems and often feel inadequately compensated [Younger *et al.*, 1977]. A recent survey of surgical ICU directors found that three out of four felt that ICU nursing was the single most constrained resource [Cullen, 1981].

preliminary findings suggest that patients displaced because of the need for a bed did not suffer adverse consequences [Williams, 1980]. These conflicting results presumably reflect the degree of rationing in different units prior to the decrease in availability.

Clinical Decisionmaking. How does the physician decide which of two patients to admit to the ICU when only one bed is available? Which patient should be displaced when a critically ill patient with acute reversible disease requires admission to an ICU that is clearly full? Generally, the practice is to allocate the bed to the patient most likely to benefit from intensive monitoring or therapy (or to displace the patient least likely to benefit).[19] But the physician's ability to estimate benefits in terms of probability of a favorable outcome is constrained by the real limits of medical knowledge. The natural histories of critical illnesses have been too imprecisely defined; the efficacy of many diagnostic and therapeutic interventions are similarly uncertain. Faced with this uncertainty, physicians often make decisions, including those that determine the allocation of intensive care, by following standard procedures. The patient with suspected myocardial infarction may be routinely kept in the CCU for three days. The patient with a rhythm disturbance may be monitored for 24 hours after the last abnormality was noted. Such broad classifications provide only the roughest estimates of risk and weighing of benefits.

Physicians allocate resources on the basis of factors other than the expected clinical course of the acute illness. Factors that influence decisions made in the management of critically ill patients with a poor prognosis with or without ICU care (Case 2) have been studied using questionnaires consisting of hypothetical case studies [Crane, 1975] and a survey of what intensive care physicians say is important for deciding the level of care for long-term ICU patients [Perkins *et al.*, 1981]. Crane hypothesized that physicians evaluate the critically ill and base treatment decisions not only on the physiological aspects of illness but also on the patient's capacity for social interaction. She distinguished between a patient's salvageability (based on the patient's prognosis) and irreversible physical or intellectual impairment that would affect his capacity to perform his established social roles. The results of this extensive study indicate that physicians consider the treatable patient one who can, or has the potential to, interact meaningfully with others. Physicians who say they would treat different types of patients ac-

[19] The criterion of maximum expected benefit does not always apply. A patient with a poor prognosis regardless of ICU care (Case 2) is more likely to be admitted and less likely to be displaced than a patient with a good prognosis regardless of ICU care (Case 3) even if the latter's expected benefit is greater. Bad outcomes for the patient with a poor prognosis may seem too directly attributable to the triaging decision.

tively, regardless of social potential, are in a small minority. Salvageable patients with intellectual impairment are less likely to be treated actively than those physically impaired.

In a smaller study, but one assessing decision factors from ICU personnel about actual patients during their ICU course, Perkins and others found similar results. Among factors given high importance by ICU decisionmakers managing patients who remained critically ill despite at least three days of intensive care were: the expected functional outcome for the patient, the patient's and family's preference regarding care, and the patient's future potential contributions to family and society. Factors said to be of little or no importance in determining care included: the patient's illness as a burden to society, the patient's past contributions to family or society, the provider's personal feelings about the patient, legal liability, and the patient's responsibility for having caused his illness. Factors said to be of intermediate importance were: hardships caused the patient by therapy, the burden of the patient's illness on the family, the educational value of treating the patient, other providers' value statements about the patient, and the dollar cost of the patient's care.

While both of these studies focus attention on critically ill patients with a low probability of survival, it is likely that these same factors influence decisions about allocation of resources among those expected to benefit from intensive monitoring because of a risk of critical illness. The patient with impaired intellectual function or with expectations for poor functional outcome may be less likely to fill the last ICU bed or be the first displaced.

The importance of functional capacity and the future orientation of ICU decisionmakers has been suggested by Crane and by Perkins and associates. This raises some questions about the rationing of intensive care by age. Excluded from some early ICUs solely on the basis of age, the elderly have become the principal consumers of intensive medical care.[20] At Massachusetts General Hospital, indications for ICU/CCU admission, major interventions, and outcomes for elderly patients have been compared with those for younger patients. Demographic comparisons were made with the population of the hospital's catchment area to allow some inferences about selective admission of the elderly [Campion *et al.*, 1981]. Of all patients admitted to the ICU/CCU at MGH, 44% were 65 years or older; nearly half of these were 75 years or older. Only 17% of adults residing in the catchment area were 65 or older. Among patients 75 or older, only 5% came from nursing homes. In the

[20] Forty-eight percent of Cullen's critically ill surgical patients [Cullen, *et al.*, 1976], 44% of the MGH ICU/CCU population [Thibault, 1980], and 26% of the GW UMC medical-surgical unit population were over 65 [Draper *et al.*, 1981].

community, nursing home residents comprised 10% of the population age 75 and over.

Once admitted to the ICU, however, the elderly were more likely to have major diagnostic and therapeutic interventions initiated during their ICU/CCU stay but were less likely to survive. Despite more frequent major interventions and higher mortality, older patients did not incur higher charges or have longer ICU stays than younger patients. This can be explained in part by the fact that elderly patients who died had lower hospital charges than did younger nonsurvivors with little differences in length of stay. There was a sharp decline with age in the ratio of mean hospital charges of nonsurvivors to those of survivors from approximately 3.0 for those age 55-64 to 1.25 for those age 75 and over. This suggests the selective allocation of resources among patients already admitted to the ICU on the basis of age. More resources were expended in the effort to prevent an inexorable decline to death in the middle-aged patient than in the older patient.[21]

The importance of prognosis in determining intensive care decisions, cited repeatedly in the clinical literature and documented by Crane and by Perkins and colleagues, begs the question: How well do clinicians predict the outcome of intensive care? Perkins and others found ICU personnel to be overly optimistic when asked to predict patients' functional status six weeks later. Predictions were made only for patients who had spent three consecutive days in the ICU. Interestingly, at the same time physicians were overestimating "the benefits," they were substantially underestimating the resources (hospital charges and length of stay) that would be expended.

At Massachusetts General Hospital, physicians were asked at the time of a patient's ICU admission to assign a probability of hospital survival and a probability of one-year survival given hospital survival [Mulley *et al.*, *Physician's Ability*, 1980]. When these probability estimates were validated against actual outcomes, it was clear that, while the probability of survival (particularly hospital survival) was generally underestimated, the estimates allowed reasonable stratification of patients into distinguishable risk groups. Short-term prognoses were more accurate than long-term prognoses. Long-term prognoses were least accurate in patients over 70 and in patients with congestive heart failure and chronic pulmonary disease.

When the prognosis estimates collected at the time of ICU admission were examined with measures of subsequent resource expenditures (total hospital charges, ancillary charges,

[21] Less severe disease among younger survivors would be an alternative explanation for the decline in the nonsurvivor-to-survivor charge ratios with age, but charges for survivors changed little with age. The decline in the ratio was due to less being spent on older nonsurvivors.

length of stay), the relationships between prognosis and expenditure among nonsurvivors and among survivors provided indirect evidence that physicians selectively allocate resources among ICU patients based on the estimates of prognosis [Detsky *et al.*, 1981]. The association between high cost and high mortality in the ICU was again confirmed; the care of nonsurvivors involved a significantly higher mean expenditure than did the care of survivors. However, among nonsurvivors, expenditure positively correlated with the probability of survival estimated at the time of admission; and, among survivors, expenditure negatively correlated with the probability of survival (Figure 3). Among both nonsurvivors and survivors, total expenditures per day were greatest for patients whose outcomes were unexpected. The nonsurvivors with poor prognoses had lower costs than survivors with similar prognoses or nonsurvivors with better prognoses.

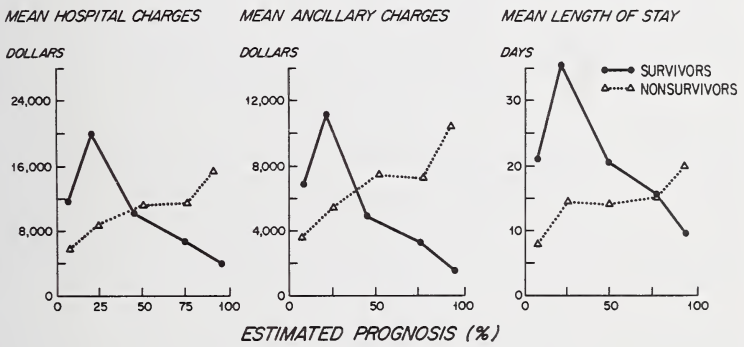


Figure 3: Relationships Between Prognosis and Resource Expenditures for Survivors and Nonsurvivors in the MGH ICU.

Source: A.S. Detsky, S.C. Stricker, A.G. Mulley, and G.E. Thibault, *Prognosis, Survival and the Expenditure of Hospital Resources for Patients in an Intensive Care Unit*, N Engl J Med 305:667-672 (1981). Reprinted with permission.

Equity, Ethics, and Policy Implications

Little can be said with assurance about the equity of current allocation of intensive care resources. Since availability determines the level of expected benefit necessary to justify intensive care, particularly for patients admitted for intensive monitoring, patients in locales with low availability may be less likely to benefit from intensive care than those in areas of high

availability.[22] Russell's finding—that intensive care beds are less likely to be available in hospitals that serve nonwhite populations and in hospitals operated by state and local governments, which are more likely to serve economically disadvantaged patients—is troublesome in this context. Because of the high costs of intensive care, those with no insurance or limited insurance are likely to mean lost revenue for hospitals when admitted to the intensive care unit. Perhaps because hospitals' bad debts rarely affect the physicians who make ICU triage decisions, there is no evidence to suggest discrimination in the allocation of ICU resources on the basis of insurance status.[23] Patients living in rural areas face access problems, and those served by small hospitals (where equal availability can be maintained only with lower average occupancy and therefore higher expenditure) may receive fewer intensive care benefits; but again, there is no direct evidence to support this.

The ethical considerations are more pressing when resources are allocated among identified, critically ill patients in the process of clinical decisionmaking. If we are to believe what physicians say they do [Crane, 1975, Perkins *et al.*, 1981], and what we can infer from empirical findings, we are led to the conclusion that physicians do discriminate in allocating intensive care resources. It appears that fewer resources are expended on the elderly, those with poorer prognoses, those with impaired functional status (particularly intellectual impairment), and those unable to fulfill their social roles. More resources are expended on those with a high probability of regaining full functional status and contributing to society.

To the extent that rationing of intensive care occurs, the practice is at odds with the generally accepted social norm that any expenditure is justified to preserve life. To the extent that it occurs on the basis of implicit estimates of prognosis, functional outcome, future contributions to society, and other subjective assessments, there is potential for discrimination. Prognostic estimates and projected functional capacity may be consciously or unconsciously influenced by the decisionmaker's affinity (or lack of it) for the patient.[24] Patients in higher

[22] This supposition must be a tentative one. Intensive care is not a standardized product. It may well be that ICU personnel who must deal frequently with major complications are more accomplished than those who care for high proportions of low-risk patients because of high bed availability. While the probability of a serious complication necessary to warrant intensive care in the low-availability unit may be greater, the probability of improved outcome given occurrence of a complication in that unit may also be greater.

[23] One would expect such evidence to be hard to come by. The potential for this form of discrimination may be increasing with the growth of the for-profit hospital sector and physician incentives for hospital cost containment.

[24] When patient variables were examined for their effect on level of

socioeconomic groups may be more likely to be cared for by physicians who are more influential in making allocation decisions. These affluent patients are more likely to benefit from their physician's influence when criteria are implicit.[25]

The implicitness and subjectivity are not without positive aspects. While some have argued that the disparity between the "sanctity of life" social norm and actual practices results in an unhealthy tension [Crane, 1975], others have stressed the value of maintaining belief in the social norm despite the practical limitations imposed by resource constraints [Zeckhauser, 1972]. Implicit criteria based on subjective assessments allow preservation of a social myth, if not the social norm.

Preserving the myth, however, cannot be justification for either inequity due to discriminatory rationing or inefficiency due to rationing in the face of ignorance. Policy must steer a course between explicit criteria that guard against discriminatory rationing and implicit judgments that soften the inevitable resources-over-life decisions. When possible, physician uncertainty regarding expected benefits should be reduced. More attention should be paid to prognoses. More precise information about the natural history of intensive care illnesses and the effectiveness of ICU interventions would at least allow more informed decisions. Simple clinical markers for patients at high and low risk for serious complications can be defined [Singer *et al.*, 1981; Mulley *et al.*, *The Course of Ssuspected Myocardial Infarction*, 1980]. Yet new knowledge about intensive care illnesses and more skilled predictions by physicians (together resulting in improved probability estimates for the decision trees in Figure 2) will not solve the ethical problems of rationing intensive care. Ethical problems will remain, because of the irreducible subjectivity of utility assessments for individual patients, and because of the tension between benefits for individual patients and societal goals that results from resource constraints. New knowledge and better prediction will, however, allow trade-offs to be more clearly defined and collective benefits derived from available resources to be more nearly maximized.

---

expenditures for ICU patients at MGH, whether or not the patient's personal physician was on the MGH staff was included as an independent variable in multiple regressions on measures of expenditures for both nonsurvivors and survivors. Having a personal physician involved in the ICU decision was associated with a shorter ICU stay for nonsurvivors but had no effect on expenditures for survivors. One interpretation is that when poor prognosis and functional impairment are well appreciated, decisions to treat less actively can be made more quickly [unpublished data].

[25] This is one reason ICUs should have a clinical director or triage officer who can make rationing decisions without being influenced by obligations to some but not other patients.

References

American Hospital Association, 1978. Hospital Statistics. Chicago: AHA, 1978.

American Hospital Association, 1980. Hospital Statistics. Chicago: AHA, 1980.

Arthur D. Little, Inc., 1979. (DHEW Publication no. (HRA) 79-14020.) Planning for Intensive Care Units: A Technical Assistance Document for Planning Agencies. Washington, D.C.: U.S.G.P.O.

Bloom, B.S., Peterson, O.L., 1973. *End Results, Costs and Productivity of Coronary-Care Units.* N Engl J Med 288:72-78.

Brown, K.W.G., MacMillian, R.L., Forbath, N., *et al.*, 1963. *Coronary Unit. An Intensive Care Center for Acute Myocardial Infarction.* Lancet 2:349-352.

Callahan, J.A., Spiekerman, R.E., Broadbent, J.C., *et al.*, 1967. *St. Mary's Hospital-Mayo Clinic Medical Intensive Care Unit II. Patient Population.* Mayo Clin Proc 42:322-338.

Campion, E.W., Mulley, A.G., Goldstein, R.L., Barnett, G.O., Thibault, G.E., 1981. *Medical Intensive Care for the Elderly: A Study of Current Utilization Patterns, Costs and Outcomes.* JAMA 246:2052-2056.

Christiansen, I., Iverson, K., Skouby, A.P., 1971. *Benefits Obtained by the Introduction of a Coronary Care Unit.* Acta Med Scand 189:285-291.

Civetta, J.M., 1973. *The Inverse Relationship Between Cost and Survival.* J Surg Res 14:265-269.

Cooper, J.K., Corcoran, T.M., 1974. *Estimating Bed Needs by Means of Queuing Theory.* N Engl J Med 291:404-405.

Crane, D., 1975. The Sanctity of Social Life: Physicians' Treatment of Critically Ill Patients. New York: Russell Sage Foundation.

Cullen, D.J., 1981. *Surgical Intensive Care: Current Perceptions and Problems.* Crit Care Med 9:263 (abstract). Presented at the Third World Congress on Intensive and Critical Care Medicine, Washington, D.C.

Cullen, D.J., Civetta, J.M., Briggs, B.A., 1974. *The Therapeutic Intervention Scoring System.* Crit Care Med 2:57.

Cullen, D.J., Ferrara, L.C., Briggs, B.A., Walker, P.F., Gilbert, J., 1976. *Survival, Hospitalization Charges and Follow-up Results in Critically Ill Patients.* N Engl J Med 294:982-987.

Day, H.W., 1963. *An Intensive Coronary Care Area.* Dis Chest 44:423-427.

Detsky, A.S., Stricker, S.C., Mulley, A.G., Thibault, G.E., 1981. *Prognosis, Survival and the Expenditure of Hospital Resources for Patients in an Intensive Care Unit.* N Engl J Med 305:667-672.

Draper, E.A., Wagner, D.P., Knaus, W.A., 1981. *The Use of Intensive Care: A Comparison of a University and Community Hospital.* Health Care Financing Rev 3:49-64.

Eisenberg, J.M., Horowitz, L.N., Busch, R., Arvan, D., Rawnsley, H., 1979. *Diagnosis of Acute Myocardial Infarction in the Emergency Room: A Prospective Assessment of Clinical Decision Making and the Usefulness of Immediate Cardiac Enzyme Determination.* J Community Health 4:190-198.

Franklin, C.M., Jackson, D.L., Koch, K.A., 1981. *Retrospective Analysis of Medical Intensive Care Unit Readmissions.* Crit Care Med 9:263 (abstract) Presented at the Third World Congress on Intensive and Critical Care Medicine, Washington, D.C.

Griner, P.F., 1972. *Treatment of Acute Pulmonary Edema: Conventional or Intensive Care?* Ann Intern Med 77:501-506.

Griner, P.F., 1973. *Medical Intensive Care in the Teaching Hospital: Costs Versus Benefits: The Need for an Assessment.* Ann Intern Med 78:581-585.

Hackett, T.P., Cassem, N.H., and Wishnie, H., 1968. *The Coronary Care Unit: An Appraisal of its Psychological Hazards.* N Engl J Med 279:1365-1370.

Harris, J.E., 1975. *The Internal Organization of Hospitals: Some Economic Implications.* Bell J Econ 6:467-481.

Harris, J.E., 1979. *Pricing Rules for Hospitals.* Bell J Econ 10:224-243.

Hill, J.D., Hampton, J.R., Mitchell, J.R.A., 1978. *A Randomized Trial of Home-versus-Hospital Management for Patients with Suspected Myocardial Infarction.* Lancet 1:837-841.

Hofvendahl, S., 1971. *Influence of Treatment in a CCU on Prognosis in Acute Myocardial Infarction.* Acta Med Scand [Suppl] 519:1-78.

Killip, T., III, Kimball, J.T., 1967. *Treatment of Myocardial Infarction in a Coronary Care Unit. A Two-Year Experience with 250 Patients.* Amer J Cardiol 20:457-464.

Klaus, A.P., Sarachek, N.S., Greenberg, D., *et al.*, 1970. *Evaluating Coronary Care Units.* Am Heart J 79:471-480.

Knaus, W.A., Draper, E., Lawrence, D.E., Wagner, D.P., Zimmerman, J.E., 1981. *Neurosurgical Admissions to the ICU: Intensive Monitoring versus Intensive Therapy.* Neurosurgery 8:438-442.

Knaus, W.A., Wagner, D.P., Draper, E.A., Lawrence, D.E., Zimmerman, J.E., 1981. *The Range of Intensive Care Services Today.* JAMA 246:2711-2716.

Knaus, W.A., Zimmerman, J.E., Wagner, D.P., Draper, E.A., Lawrence, D.E., 1981. *APACHE—Acute Physiology and Chronic Health Evaluation: A Physiologically Based Classification System.* Crit Care Med 9:591-597.

Lepore, M.J., Grace, W.J., 1969. *Role of the Intensive Care Unit in Gastroenterology.* Am J Gastroenterol 51:493-497.

Lown, B., Fakhro, A.M., Hood, W.B., Jr., and Thorn, G.W., 1967. *The Coronary Care Unit. New Perspectives and Directions.* JAMA 199:188-198.

Martin, S.P., Donaldson, M.C., *et al.*, 1974. *Inputs Into Coronary Care During 30 Years: A Cost-Effective Study.* Ann Intern Med 81:289-293.

Mass. Physician, 1975. *Intensive Care: How Valuable?* 34 (7):15-16.

Mather, H.D., Morgan, D.C., Pearson, N.G., Read, K.L.Q., *et al.*, 1976. *Myocardial Infarction: A Comparison Between Home and Hospital Care for Patients.* Br Med J 1:925-929.

Morgan, A., Daley, C., Murawski, B.J., 1973. *Dollar and Human Cost of Intensive Care.* J Surg Res 14:441.

Mulley, A.G., 1980. *Intensive Care Decision Making.* Presented at the Annual Meeting of the Society for Medical Decision Making, Washington, D.C.

Mulley, A.G., Thibault, G.E., Barnett, G.O., *et al.*, 1980. *Physicians' Ability to Predict Outcome of Medical Intensive Care.* Clin Res 28:301A (Presented at the Annual Meeting of the American Federation for Clinical Research, Washington, D.C., May).

Mulley, A.G., Thibault, G.E., Hughes, R.A.,. Barnett, G.O., Reder, V.A., Sherman, E.S., 1980. *The Course of Suspected Myocardial Infarction: The Identification of Low Risk Patients for Early Transfer.* N Engl J Med 302:943-948.

Oliver, M.F., Julian, D.G., Donald, K.W., 1967. *Problems in Evaluating Coronary Care Units. Their Responsibilities and Their Relation to the Community.* Am J Cardiol 20:465-474.

Perkins, H.S., Jonsen, A.R., Epstein, W.V., 1981. *The Importance and Accuracy of Treatment Decisions about Long-Term Intensive Care Patients.* Submitted for publication.

Piper, K.W., Griner, P.F., 1974. *Suicide Attempts with Drug Overdose: Outcomes of Intensive vs. Conventional Floor Care.* Arch Int Med 134:703-706.

Pitner, S.E.,. Mance, C.J., 1973. *An Evaluation of Stroke Intensive Care: Results in a Municipal Hospital.* Stroke 4:737-741.

Pontoppidan, H., Abbott, W.M., Brewster, D.C., *et al.*, 1976. *Options in Care for Hopelessly Ill Patients.* N Engl J Med 295:362-364.

Redisch, M.A., 1978. *Physician Involvement in Hospital Decision Making.* In Hospital Cost Containment, edited by M. Zubkoff, L.E. Raskin, R.S. Hanft. New York: Prodist.

Rogers, R.M., Weiler, C., Ruppenthal, B., 1972. *Impact of the Respiratory Intensive Care Unit on Survival of Patients with Acute Respiratory Failure.* Chest 62:94-97.

Russell, L.B., 1979. Technology in Hospitals: Medical Advances and Their Diffusion. The Brookings Institution, Washington, D.C.

Scitovsky, A.A., 1967. *Changes in the Costs of Treatment of Selected Illnesses, 1951-1965.* Am Econ Rev 57:1182-1195.

Shepard, D.S., Ghanotakis, A.J., 1979. *Hospital Costs in Massachusetts: A Report of the Massachusetts Funds Flow Project.* Springfield, Va.: National Technical Information Service (HRP-0029335).

Singer, D.E., Mulley, A.G., Thibault, G.E., Barnett, G.O., 1981. *Unexpected Readmissions to the Coronary Care Unit During Recovery from Acute Myocardial Infarction.* N Engl J Med 304:625-629.

Steel, K., Gertman, P.M., Crescenzi, B.S., *et al.*, 1981. *Iatrogenic Illness on a General Medical Service at a University Hospital.* N Engl J Med 304:638-642.

Tagge, G.F., Adler, D., Bryan-Brown, W., Shoemaker, W.C., 1974. *Relationship of Therapy to Prognosis in Critically Ill Patients.* Crit Care Med 2:61-63.

Thibault, G.E.,. Mulley, A.G., Barnett, G.O., *et al.*, 1980. *Medical Intensive Care: Indications, Interventions, and Outcomes.* N Engl J Med 302:938-942.

Turnbull, A.D., Carlon, G., Baron, R., Sichel, W., Yang, C., Howland, W., 1979. *The Inverse Relationship Between Cost and Survival in the Critically Ill Cancer Patient.* Crit Care Med 7:20-23.

Virginia Regional Medical Program Cardiac Study Group, 1975. *An Analysis of Cardiac Care Unit Experiences.* Virginia Med Monthly 102:205-213.

Wagner, D.P., 1980a. *National Costs of Intensive Care.* Presented at the Symposium, Critical Issues in Medical Technology, Boston.

Wagner, D.P., 1980b. *The Demographics of George Washington University's Intensive Care Unit.* Presented at the Symposium, Critical Issues in Medical Technology, Boston.

Williams, S.V., 1980. *The Consequences of Premature Discharge from the Medical Intensive Care Unit.* Presented at the Symposium, Critical Issues in Medical Technology, Boston.

Younger, S., Jackson, D.L., Allen, M., 1977. *Staff Attitudes Towards the Care of the Critically Ill in the Medical Intensive Care Unit.* Crit Care Med 7:35-40.

Zeckhauser, R., 1972. *Coverage for Catastrophic Illness.* 1972 Discussion Paper. Public Policy Program, Kennedy School of Government, Harvard University.

Acknowledgements

I am indebted to Dr. William Knaus, Douglas Wagner and Elizabeth Draper of the George Washington University ICU Research Unit for sharing their data and insights; to Dr. Henry Perkins of the Institute for Health Policy Studies at the University of California, San Francisco for providing information about physicians' rationing criteria and predictive ability; to Drs. Daniel Singer and George Thibault of the Medical Practices Evaluation Unit and Drs. Michael Rie and David Cullen of the Department of Anesthesia, Massachusetts General Hospital for their helpful comments and criticisms; and to Carol Scola for her assistance in the preparation and editing of the manuscript.

Health Maintenance Organizations and the Rationing of Medical Care

X

Harold S. Luft*

HEALTH MAINTENANCE ORGANIZATIONS (HMOs), with their capitation payments and contractual obligations to provide services, present a set of economic incentives that depart significantly from those of the conventional fee-for-service medical care system. It might be expected that these incentive differences will result in different approaches to the allocation of medical resources. Furthermore, these allocation or rationing mechanisms might have ethical implications because of the differential impact on various socioeconomic groups.

This paper will address a series of questions arising from the differences between prepaid and fee-for-service systems. The first section will examine the incentives under each system. In doing so, it will focus on incentives to the enrollee or patient, to the organization, and to providers within the system. This discussion will highlight the conflicting incentives inherent in various arrangements. The second section will examine how these incentives and conflicts may be dealt with in terms of the rationing or allocation of resources. The third section will expand the analysis to consider the probable effects of different systems with respect to various socioeconomic groups. This discussion will focus on both the microquestion of the impact of various mechanisms to allocate services among plan members, and the macroquestion of the effects of HMO enrollment and marketing decisions on who gets access to HMO options. To the extent possible, this discussion will include empirical evidence on the magnitude of

This study, which was originally prepared for the Commission, is reprinted with permission from MILBANK MEMORIAL FUND QUARTERLY/ HEALTH AND SOCIETY, Vol. 60, No. 2, 1982.

*Institute for Health Policy Studies, University of California, San Francisco.

such effects. The fourth section extends the discussion to speculate on the possible effects of an increasingly competitive environment for HMOs and other health care plans. Again, while this section will be based largely on conjecture, some evidence is available for analysis. The final section of the paper will consider the ethical implications of the issues already raised and will discuss some of the policy implications arising from them.

Economic Incentives in Various Medical Settings

Economic incentives and allocation mechanisms exist in all medical care delivery systems. Defining either optimally efficient or ethically just systems within the context of medical care realities is a task well beyond the scope of this paper. Instead, the focus here is on how various prepaid systems, of which HMOs are one notable example, *differ* from the conventional, primarily fee-for-service system. Subsequent discussions may then address the desirability of alternative systems.

Incentives in the Conventional Fee-for-Service Setting

The primary actors in the medical care system are: 1) consumers or, after they begin to receive medical care, patients; 2) insurance companies or other third parties; 3) physicians; and 4) hospitals. For most employed people and their families, health insurance is purchased through employer or union groups, often with a substantial employer contribution. The insurance plan or plans to be offered are selected by the employer or union and in most cases only one option is available to the employee. When ill, the patient selects a physician to diagnose the problem, provide treatment, and serve as an agent in arranging for the provision of services by hospitals and other physicians. All of these providers are typically paid on a fee-for-service or cost-reimbursement basis by the patient, the patient's insurer, or the government.

To the extent that insurance or other third-party payments such as Medicaid reduce the net price of services to the patient, more medical care will be demanded than if the patient paid the total bill (Newhouse et al., 1981). How much greater demand will be, or in economic terms the price elasticity of demand, is a much debated empirical question, but it is generally agreed that there is some increase, and that it is proportionately greater for discretionary and ambulatory services than for emergency and hospital services. Simultaneously, fee-for-service payments to the providers create a natural incentive to do more and respond to the increased demand.

Incentives in an HMO Setting

An HMO changes some of these linkages in ways that alter substantially the economic incentives. The patient no longer pays fee-for-

service, except for an occasional small copayment of a dollar or two, so demand is likely to increase relative to conventional coverage. The HMO can also develop a set of financial linkages that tie together the physician, hospital, and insurer in a fashion quite unlike that of the conventional system. In contrast to the typical fee-for-service situation, the physicians and hospitals gain economically by doing less. This sets up a direct conflict with their patients whose comprehensive coverage leads them to demand even more than is usually the case.

In HMOs where the physicians practice as a group, two payment schemes are generally used: 1) salaries, with some type of bonuses if the group performs well; or 2) fee-for-service based payments modified to reflect the total pool of funds available. Salaries, of course, provide an incentive to exert oneself as little as possible unless there are other rewards for good performance. Fee-for-service payments tend to increase productivity—as measured by services per hour and hours worked per year—but increase the conflict between incentives to do more and the group incentives to do less (Held and Reinhardt, 1980).

In addition to the group model HMO, there is the individual practice association (IPA) model which typically involves an HMO acting as a health plan which then contracts through an intermediary organization with independent office-based physicians to provide services. These physicians agree to accept payment by the HMO as payment in full and to allow a withholding of these fees to be placed at risk. The withheld fee is paid back only if the HMO meets its financial goals. Thus, while the individual physician can still earn more by performing more services, he or she can lose money if the physicians collectively provide or order too much or too expensive care.

The implications of these different economic incentives on the supply of services by the HMO and its providers are quite complex. When the physicians are essentially salaried and share in the financial success of the plan, they will collectively and individually probably try to reduce services. When physicians are essentially paid on a fee-for-service basis within a group or IPA type HMO, this incentive to do more will be only partially offset by the collective risk-sharing. This is especially true in IPAs in which the prepaid patients typically account for under 15 percent of the physicians' caseload.

Organizational Influences

While economic incentives are clearly different in HMOs, one must also consider a second dimension of difference—the organizational setting. The prepaid-group-practice type of HMO, which has been the economically most successful model, is characterized by both

prepayment *and* the group-practice mode of delivery. Often, one of the attractions for physicians of large group practices is a more regularly scheduled life, with vacation time, educational leave, shorter hours, and on-call coverage. Such arrangements imply less doctor-patient continuity and typically a more bureaucratic organization (Luft, 1981; Mechanic, 1976). Prepaid groups, in particular, seem to have developed appointment and scheduling systems that reduce patient waiting time in the office but increase the time one must wait to obtain an appointment. Patients who drop in or request a visit the same day are typically seen in an urgent-care clinic, usually by a physician on rotation, rather than their usual provider.

In the discussion that follows it will not be possible to examine all possible combinations of economic and organizational factors that may influence the provision of services. Instead, three models of practice will be discussed. The first is the conventional system of independently practicing fee-for-service providers with third-party insurance available for the patient. The second is the individual practice association or IPA model HMO involving primarily independent fee-for-service practitioners linked through a risk-sharing arrangement with an HMO. The third is the prepaid-group-practice or PGP model HMO composed of a large multispecialty-physician group practice which obtains most of its revenue from the HMO in the form of fixed capitation or salary payments plus some risk sharing based on overall plan performance.

Methods of Allocating Services in HMOs

In classical market situations, goods and services are allocated through a price system. In medical care, various factors combine to cause substantial deviations from a market in which price is the only means of allocation. Information is scarce, both because of restrictions on advertising and, more importantly, because of the technical nature of the product, so consumers cannot make rational decisions (Arrow, 1963). Not only must the patient rely upon the physician to diagnose and provide services, but in many instances it is difficult to check on performance, so trust becomes an important factor. Medical services also require the presence of the patient and thus impose a time cost to obtain medical care and information about treatment alternatives.

Further complicating the analysis of the allocation of medical services is the fact that two levels of decisions are involved. The first decision is the type of insurance *cum* medical delivery system, that is, conventional insurance with independent fee-for-service providers or a prepaid system with a more limited choice of providers. (Of course, one also must consider *which* of the available providers one will choose within these two groupings.) The second level of decision

concerns how much medical care will be consumed in treating a specific problem. Each type of decision reflects the preferences and behaviors of consumers/patients, on the one hand, and insurers/providers on the other. HMOs and insurers can attempt to influence the enrollment choices of consumers, and providers can influence or control the services delivered. Consumers often have strong preferences for practice setting (group vs. solo) and constrained vs. free choice of provider. Patients also have preferences, sometimes strongly held ones, as to the treatments they expect for specific illnesses. The following discussion will examine first the factors that influence treatment decisions and how these might vary among different practice settings; these might be termed "service allocation or rationing" issues. The focus will then widen to consider the allocation of consumers among systems; these might be termed "enrollment allocation or rationing" issues.

Decisionmaking in Three Types of Medical Treatment

When examining how medical care is allocated under various systems, it is useful to distinguish three major types of services: 1) patient initiated visits; 2) physician initiated or controlled visits; and 3) hospitalizations (Gertman, 1974). In a typical situation, a person recognizes certain problems or symptoms and decides that an examination by a physician is warranted. Obviously, whether a physician is actually seen will depend on the fee that must be paid, time costs, transportation and access costs, and whether the patient knows a potential provider. At the end of the examination, the physician may request a follow-up visit or may suggest a consultation with a specialist. (Of course, this is not entirely up to the physician; the patient may request such follow-up procedures for more reassurance and, likewise, the patient may refuse to follow through.) In some cases, the patient may continue to initiate follow-up visits by seeking out physicians. In general, the consumer has less control with respect to hospital services—admission can be gained only with the authorization of a physician and once in the hospital, the physician essentially controls what services will and will not be provided.

A crucial point in this discussion is the idea that not only do physicians have a great deal of discretion in treatment and referral decisions, but that wide variation in practice patterns may be consistent with good quality care rendered in a careful and professional manner. Some physicians will have broad indications for surgery while others will impose stringent criteria for even suspected appendicitis. Some will make aggressive use of laboratory tests to rule out conditions while others will rely more upon the patient's history and physical

examination. In most instances, large scale trials have not been performed to identify the course of action leading to the best results. Wide variations in practice patterns often coexist in similar practice settings (e.g., among independent fee-for-service practitioners in the same state), with no noticeable differences in outcomes (Roos and Roos, 1981; Wennberg and Gittelsohn, 1973; Wennberg et al., 1980). This suggests that physicians in various HMOs and fee-for-service settings may consistently differ in their recommendations for treatment, yet still be within acceptable norms.

Likely Impact on Services in Different Settings

The likely impacts of the differences in price, accessibility, and practice styles on utilization of services in different settings are summarized in Table 1. While precise measurements of these effects are unavailable, in most instances each aspect can be identified as either increasing or decreasing utilization. One must note, however, that it is not always clear whether increased or decreased utilization is necessarily desirable. In general, increased access for patient-initiated visits is probably desirable, but one can even question that assumption. For instance, annual physical examinations for certain population groups are not cost-effective (Collen et al., 1973). Moreover, frequent testing of healthy people increases the likelihood of false-positive results and subsequent anxiety, invasive testing, and iatrogenesis.

The various factors listed in Table 1 influence utilization through different mechanisms. Some, like net price, travel time and costs, and waiting time, enter into the consumer's implicit calculation concerning whether the perceived problem is worth a visit. In general, because IPA-HMOs involve physicians primarily in fee-for-service office based practice, the only difference perceived by their HMO patients relative to their fee-for-service patients is more comprehensive coverage. The resulting lower net price will tend to increase utilization. Prepaid group-practice (PGP) patients also face a lower (or zero) net money price, but other factors change simultaneously. They may not have a close relationship to a physician in the plan and thus may find it more difficult to take the initial steps to obtain treatment. PGPs are typically centralized so that for most patients a longer trip will be involved. The PGP also generally involves a different set of time costs—appointments must be scheduled further in advance, but once made, waiting time in the office is typically less (Mechanic, 1976; Richardson et al., 1977; Luft, 1981). Thus, while fees are not used by HMOs as a rationing device to influence demand, as is the case in the conventional system, there are other changes associated with the organization of group practice that make access both easier and more difficult. More importantly, as will be seen below, some

of these influences on access may have different effects for people in different socioeconomic groups.

As one examines physician-initiated referral and follow-up visits and hospitalizations, differential physician behavior becomes more important relative to consumer factors. In part this is because conventional insurance policies provide better coverage for hospital services, tests, and procedures than for initial office visits, and because follow-up visits are more likely to be beyond the initial deductible. The greater physician influence also reflects the fact that when evaluating whether to make an initial visit, patients have only their own perceptions (as well as those of friends and family) concerning the need for a visit. For follow-ups, referrals, and hospitalizations, the physician's presentation of the severity of the problem will play a key role. Thus, differential clinical decision rules are a potentially major factor in the HMO's control over cost and utilization. This role is highlighted when one recognizes that in general, HMO enrollees are somewhat more likely than people in conventional plans to see a physician at least once a year, yet among utilizers, HMO enrollees have somewhat fewer office visits and substantially fewer hospitalizations (Luft, 1978).

Thus far, the discussion may have suggested that the capitation or risk-sharing incentives of the HMO lead HMO physicians to alter their practice patterns for their members. In fact, it is extremely unlikely that HMO physicians reflect upon the impact on their bonuses each time they consider a follow-up visit or an extra test. Instead, certain routine patterns are probably developed that tend to be consistent with their economic incentives. Inconsistent patterns may be reexamined and slowly adjusted to reduce conflict with system incentives.

As before, the situation in IPA and PGP HMOs differs markedly. While the research has yet to be done, it seems unlikely that many physicians will use different decision rules for their fee-for-service and prepaid patients. The available evidence supports the notion that current levels of risk sharing and patient loads in IPAs (10–20 percent withheld fees and under 15 percent of the patients) have little impact on performance (Meier and Tillotson, 1978). Instead, IPAs exercise controls through hospital-based utilization review and through educational efforts to change general practice patterns by showing that certain less costly techniques are equally effective.

PGP physicians typically see only HMO patients, so the altering of practice styles may be easier because a "two-class" system is not necessary. Some HMOs develop extensive data bases on the use of tests by physicians and feed back this information. Merely seeing that he or she has an ordering pattern very different from the rest of the group is often sufficient incentive for a clinician to rethink decision

TABLE 1
Probable Effects on Utilization of Services of Various Factors in Different Practice Settings

Type of Medical Care/ Rationing Factors	Conventional Fee-for Service and Insurance	IPA-HMOs: Fee-for-Service Practitioners at Risk	PGP-HMOs
Patient Initiated Visits			
Price to consumer	Initial and preventive visits often not covered −	Comprehensive coverage of all visits +	Comprehensive coverage of all visits +
Knowledge of provider	Often a local physician with a longstanding relationship +	Often a local physician with a longstanding relationship +	Often a local physician with no prior contact with patient −
Appointment lag	Typically short, urgent visits "squeezed in" +	Typically short, urgent visits "squeezed in" +	Typically long, urgent visits routed to separate clinic −
Accessibility to provider	Decentralized, likely one close to patient +	Decentralized, likely one close to patient +	Centralized, generally further from patient −
Waiting time in office	Variable, often long because patients "squeezed in" −	Variable, often long because patients "squeezed in" −	Typically short if appointment made in advance +
Physician Initiated Visits			
Physician incentives	Follow-up increases revenue +	Follow-up increases revenue more than risk sharing +/−	Follow-up reduces net income; substitute call-backs −
Other factors similar to above			

Physician Initiated Referral			
Physician incentives	Reciprocal referrals among different specialists encouraged by professional network, discouraged by prohibitions on fee splitting +	Referrals encouraged by professional network but discouraged by risk sharing +/−	Referral attractive to "dump" a problem patient, but collegial and financial costs if frequent −
Price to consumer	More likely covered than initial visit, still not complete +/−	Comprehensive coverage of all visits +	Comprehensive coverage of all visits +
Accessibility of provider	Typically at a different location −	Typically at a different location −	Centralized—"one stop care" +
Incentives to "return patient to primary care physician"	Depends on nature of referral network +	Depends on referral network, sometimes encouraged by HMO +/−	Typically encouraged by the system −
Hospitalization			
Price to patient	Often fairly comprehensive coverage but some copayments −	Comprehensive coverage pays in full +	Comprehensive coverage pays in full +
Incentives for physician	Hourly income higher in hospital +	Hourly income higher, but risk sharing tends to counter +/−	No additional income, costs are borne by plan −

\+ = tends to increase utilization
− = tends to decrease utilization
+/− = mixed effects

rules. In other cases, active physician-education programs have been undertaken by the HMO (Thompson, 1979). Another, perhaps crucial factor is self-selection by physicians. PGPs are often perceived by outside physicians as imposing restrictions on the practices of their clinicians and this is likely to deter even initial investigation by physicians who value highly the unconstrained ability to aggressively use any treatment. (Interestingly, PGP physicians often mention as one of the advantages of prepaid practice the ability to select treatments without having to worry about the financial cost to the patient [Cook, 1971].) Similarly, physicians who value a wait-and-see approach may be attracted by the HMO setting which does not penalize them financially for not ordering batteries of tests. In this way prepaid and fee-for-service settings may attract physicians with previously developed practice styles consistent with their subsequent economic incentives.

Allocation of Consumers Among Systems

Thus far, the focus has been on the provision of services within an HMO or other medical care setting. However, it is important to also take a broader view and examine the question of who is allowed into the setting. HMOs, like other insurers, receive most of their members through employee groups. Generally only a small number of people are allowed to enroll as individual (nongroup) contracts and their premiums are typically higher. (In fact, the record of HMOs on this issue is generally better than that of most commercial insurers. The HMO Act of 1973 requires federally qualified HMOs to provide an open enrollment season and allow individual subscribers to enroll at community rates. Conventional insurers are under no such obligation.)

If an HMO wishes to select a certain type of enrollee—healthy, white, female, upper income, or whatever—two strategies may be used. The first involves careful selection of the employee groups to whom the HMO option will be offered. For instance, firms with primarily white collar workers might be placed high on the marketing list while other firms might be ranked lower. (The mandating provisions of the HMO Act allow a federally qualified HMO to require certain employers to offer the HMO as an option. There is no parallel requirement of an HMO to offer enrollment to any employer requesting coverage.) In general, it is difficult to document instances in which firms were not approached because their employees were thought to be too risky—numerous explanations for marketing patterns can be described. However, there are some examples of behavior toward firms already in the HMO that support the notion of these techniques. In one case, an HMO in Minneapolis-St. Paul dropped its contract with a large local employer because maternity use within the group was too high (Matlock, 1980). In another case, an HMO

that phased in mental health benefits offered them first to working class groups least likely to use them while delaying coverage of university groups anticipated to be higher users.

The second strategy for selection would be to attract certain types of people *within* the employee group. Various tactics can be used to influence enrollment. Premiums are often weighted so that family coverage is proportionately less expensive than individual coverage, thus attracting families with children who tend to cost less per capita than single persons or two person families. (This strategy is also used by conventional insurers.) HMOs have traditionally offered more comprehensive coverage of maternity care, again attracting young, relatively healthy families. (Note that not all selection strategies are socially undesirable.) In other instances, the same factors that influence the use of services by members also influence who becomes a member. The single most important determinant of whether someone will choose a prepaid group practice when given the option is whether that decision will break a strong existing tie to a physician; people with strong ties do not join, whereas those without such ties are much more likely to join prepaid groups (Berki and Ashcraft, 1980). One explanation for the absence of a strong physician-patient relationship is low utilization of services. By chance or design, the PGP structure both attracts people without strong physician bonds and does not foster their development—and both features may lead to lower use of services (Scitovsky et al., 1979).

Another approach is in the careful location of facilities. While enrollment may be offered to any employee within a relatively wide region, if the clinics are only convenient to certain population subgroups, then enrollment is likely to reflect those locational factors. Similarly, access might be differentially altered for certain types of patients. For instance, appointment scheduling might be made easy and flexible for pediatrics and difficult for adult or geriatric medicine. The dual-choice option in the HMO is, in part, designed so that dissatisfied enrollees can leave the plan, removing both the source of complaint and the pressure on the organization to change (Hirschman, 1970; Phelan et al., 1970).

Because the typical IPA-model HMO utilizes the practices of a large number of independent practitioners, it has less control than the PGP over certain features of its delivery system, but it has more control over others. In particular, IPAs occasionally drop physicians from their roster because of excessive utilization. The situation may not have to get to the point of formal severance. Repeated review and questioning of a physician's treatment patterns is probably sufficient to cause withdrawal.

While some of the factors that have been discussed are rather subtle, they need only to result in relatively small differences among certain

people to have a major cost impact from the perspective of the HMO. In general, a small fraction of the plan's enrollees account for the majority of its costs. Most people are fairly healthy and use ambulatory care and occasional hospital services. Clearly, it is in the plan's interests if certain features, whether by design or accident, tend to keep out those people who are frequent users of costly services. Note that not only does the plan benefit, but the relatively low-cost people who do enroll receive comprehensive coverage and usually excellent service at lower average cost, at least in part because they do not have to share the costs of their fellows. Conventional insurers have the same incentives to select healthy enrollees.

This leads to a consideration of the environment in which HMOs exist. The notion of risk spreading through community rating whereby everyone in an area pays the same premium has been almost universally replaced by experience rating whereby firms with young, healthy employees pay lower premiums while those with older and more costly employees have higher premiums. In spite of this, HMOs were traditionally community rated long before the HMO Act required them to do so. Only recently, in the Omnibus Budget Reconciliation Act of 1981, P.L. 97-35, has the law been changed to allow HMOs to take into account differences in enrollment mix. How they will use the increased flexibility will depend in part on future changes in the competitive environment.

Socioeconomic Differences in the Effects of Alternative Allocation Systems

The major feature of the HMO system from the consumer's perspective is that prices in the form of coinsurance and deductibles are absent. This suggests immediately that income and ability to pay will not influence utilization, so the distribution of care may be expected to be more equitable. (In this case, equitable is taken to mean that services are allocated according to need rather than effective demand.) However, the previous section identified other variables that may influence utilization and these may have differential effects by income group, race, sex, residence, or other factors.

To economists, the most familiar device used by HMOs to constrain patient demand is time. Waiting time in an office is "dead," unproductive time. Many economists value time at the individual's wage rate so that an hour for someone earning $10 per hour is worth four times that of someone earning $2.50 per hour. Other economists would argue that wage rates alone are not appropriate measures of the value of time to an individual. At a minimum, wage rates must be compared to total income. Furthermore, a highly paid professional

typically has more sick leave, flexibility in hours, and discretion than a dishwasher paid the minimum wage. For the dishwasher, a long appointment wait may mean peremptory firing, not just an inconvenience. In any case, it is difficult to determine the relative time costs in HMOs. Travel time is generally increased in PGPs, but waiting time in the office is reduced. Furthermore, patients with appointments probably experience less variability in waiting time in PGPs. While it is difficult to know whether the shorter office waits benefit more the higher or lower income groups, they may make less difference to the retired who are probably not pressed for time, and there are numerous anecdotes of the waiting room as a social hall.

The long appointment lag in PGPs has a different type of effect. It has little impact on people who have postponable, schedulable problems with a low anxiety component. Middle-class professionals are likely to fall into this category. The poor, parents of small children, and those elderly persons with recurrent urgent or acute problems may not be served as well by such a system. However, most PGPs offer 24-hour emergency coverage and urgent drop-in visits at the clinic, so services are available, albeit with longer waits and usually an unfamiliar provider. The PGP does have a substantial advantage over the typical emergency room in that the patient's medical record is usually available.

Physical access to PGP HMOs may have a differential effect on various subgroups. The centralized location means that the poor and others without automobiles may find access more difficult. However, the presence of consultants, laboratory, X-ray, and pharmacy at the same site may be a special convenience to such people. Like institutions, large PGPs are also more likely to enhance access for certain subgroups, such as non-English-speaking minorities and the disabled. For instance, Kaiser offers a special teletypewriter (TTY) telephone line for hearing-impaired persons, a convenience not offered by many independent physicians.

The large, bureaucratic structure of some PGPs may create barriers for the less educated. There is often a "system" that one must learn how to "work," rather than the helpful receptionist in an independent practitioner's office. On the other hand, the conventional system often offers no assistance in getting from one provider to another. The HMO also eliminates the need for complicated claim forms which often increase substantially the cost to the poor and uneducated because they do not claim the benefits due them.

Prepaid group practices are by their nature closer to some people than others. This may lead to much higher travel costs for residents of some neighborhoods. In most instances, because HMOs concentrate on enrolling employed populations, this implies reduced accessibility for people in lower-class areas. However, some HMOs have targeted

their services to Medicaid recipients and thus are inaccessible to the middle and upper class.

HMOs can selectively try to enroll certain population subgroups. Clearly, they have incentives to avoid those groups with higher-than-average costs and utilization unless their premiums are commensurately higher. (Note that, in part, this incentive is caused by the HMO's community rating while other insurers rate by experience. An experience-rating HMO might even find high-risk people especially desirable because of the larger potential savings it could offer relative to the conventional system.)

Enrollment Choices by HMOs and the Poor

The most important factors in HMO selection with respect to the aged and the poor are policies set by the federal and state governments. Current Medicare provisions do not make it attractive for HMOs to enroll the elderly, nor are Medicare beneficiaries given any incentives to enroll in HMOs (Luft et al., 1980). The Medicaid program allows states to contract with HMOs to provide care for the poor. In some states this has resulted in rather large programs, but in many areas Medicaid enrollment is minimal even though HMOs are available.

The poor, almost by definition, enter HMOs under circumstances different than those of the middle class who have a dual-choice option through their employers or unions. In general, the conventional health-insurance alternative for employees provides fairly comprehensive coverage with some copayments and deductibles. Moreover, the middle-class family can generally use their conventional health insurance at a wide range of medical care providers. For those who join, the HMO is the better of several fairly good options.

Poor enrollees in HMOs have generally entered through either of two routes—a demonstration project or a Medicaid contract. In a demonstration project or experiment, a government or private agency offers to pay the premiums for those people who enroll in the HMO. In a few instances, the family is given a choice of plans, but in almost all of these instances special monitoring programs and evaluations are included. (Such efforts are likely to influence, as well as measure, performance.) These demonstration projects often include special features such as outreach coordinators and transportation programs not normally provided by the HMO (Freeborn et al., 1980; Coltin et al., 1978; Richardson et al., 1977). In some cases the experiment provided comprehensive coverage as long as the family stayed enrolled, but disenrollment would imply a loss of benefits. It is also the case that most demonstration projects for which data are available were undertaken by mature HMOs with large stable enrollments of employed

populations. Many of the HMOs supported substantial research staffs and most undertook the demonstration at least in part out of perceived social responsibility. All of these factors render such situations somewhat atypical and may alter the actual or perceived performance of the plan relative to a simple competitive market model.

In several states HMO enrollment has been made available as an option to Medicaid beneficiaries. Although on the surface such arrangements appear analogous to the multiple-choice options available to employees, there are also some important differences. One is that conventional Medicaid programs generally do not include copayments or a premium paid by the enrollee, nor is the HMO able to offer additional benefits to the enrollee should it realize savings. Thus, the Medicaid recipient is not attracted to the HMO by its better financial coverage, lower cost, or improved benefits, as is the case for middle- and working-class enrollees. Instead, the reason many Medicaid beneficiaries take an HMO option is that few fee-for-service physicians in their area are willing to accept Medicaid patients. Rather than being attracted by a better system, the HMO may offer the only viable system and thus be under less pressure to perform well. In addition to the enrollees' incentives, it is important to examine the motives of the HMO in negotiating a Medicaid contract. In most instances, this process is long and difficult, and the capitation rates offered by the state are often unattractive relative to premiums paid by employee groups. Thus, some plans that choose to attract Medicaid beneficiaries may do so because of their own precarious position in the market. This is certainly not always the case, however, and some well-known and highly successful HMOs have tried to enroll the poor out of a sense of community.

Service Allocation Differences for Different Socioeconomic Groups

One problem in searching for differential consequences by socioeconomic group is choosing the appropriate comparison population. One can argue that the poor have more health problems than the nonpoor, so the poor in an HMO should be compared to the poor in the traditional system. Alternatively, one might reason that the fee-for-service system is not now meeting the needs of the poor. Thus the experience of the poor in an HMO should be compared to that of the middle class in the same organization. Studies of both types are available.

Most of the findings concerning utilization of comparable groups of the poor in HMOs and fee-for-service settings are discussed in Luft, 1981 (see also Vignola and Strumpf, 1980). The major studies are those of Gaus, Cooper, and Hirschman (1976) on ten HMOs; Fuller

and Patera (1976) on Group Health Association in Washington, D.C.; Richardson et al. (1977) on the Seattle Model Cities Project; Johnson and Azevedo (1979) on Kaiser-Portland; and Salkever, German, Shapiro, Horky, and Skinner (1976) on East Baltimore Health Plan. For ambulatory care, the results are mixed; in half, there are more visits by the poor in HMOs, and in half, the results are reversed. These average findings are misleading because it is much more frequently the case that the poor in HMOs have at least one visit more than the poor seeing fee-for-service physicians. This suggests that access is easier for the poor in HMOs than for those in the conventional systems, even with Medicaid coverage.

For hospitalization, the results are remarkably consistent—lower admission rates for HMO enrollees in all cases. The prevailing opinion is that this generally is not due to an inability of patients to gain admission. In fact, somewhat lower utilization is probably good because of the presumed excessive hospitalization in conventional settings. While exceedingly low utilization rates may not be good practice, more detailed studies are necessary in order for us to determine if this is the case.

The second type of comparison involves the poor and nonpoor in the same organization. The question is: Does enrollment of the poor in the same plan with the middle class, and with the same financial coverage, eliminate the general underservice of the poor found in the conventional system? In general, ambulatory utilization rates of the poor in mature HMOs approximates that of the middle class.

However, some differences appear in the *types* of utilization by different socioeconomic groups. The poor are more likely than the middle-class enrollees to walk in or use the hospital emergency room instead of a regularly scheduled appointment. The former results in fragmented care, but may better fit the life circumstances of the poor. Further adding to the fragmentation was the fact that when appointments were made, the poor had a higher no-show rate (Hurtado et al., 1973; Richardson et al., 1977; Group Health Cooperative of Puget Sound, 1970; Shragg et al., 1973). It is difficult to identify the cause of these utilization patterns, but it appears that to some degree the style of service provided by the HMOS does not match that desired by the poor enrollees. Of course, the life circumstances of the poor may lead them to favor drop-in visits, emergency room use, and cancelled appointments in all settings, not just in HMOs.

These studies examined the behavior of special subgroups of the poor who were introduced to an ongoing HMO as part of a demonstration project. Such an arrangement requires adjustments by both the HMO staff and the new enrollees. An alternative research design examines whether different utilization patterns occur by socioeconomic group within the regularly enrolled population. While this approach

is limited to employed people with an HMO option (and thus omits the very poor), it provides better measures of long-run equilibrium. Furthermore, if differences appear, they are likely to be larger if coverage is extended to the very poor.

Much of this work has been done by the Kaiser Portland Research Center. Freeborn, Pope, Davis, and Mullooly (1977) summarize their results concerning outpatient utilization with respect to socioeconomic status. Health status measures appear to be the dominant factors, and once they are held constant, education, income, and socioeconomic class have little effect on utilization. The one exception is preventive service, which is not related to health status but, for women, is positively related to education and income. These studies also exhibit differences in the patterns of ambulatory use that are related to socioeconomic status—the middle class is more likely to use the telephone for reporting symptoms and is less likely to use walk-in clinics (Pope et al., 1971; Weiss and Greenlick, 1970; Nolan et al., 1967).

Hetherington, Hopkins, and Roemer (1975) argue that in PGPs people with more education and higher incomes can more easily "work the system." While it is true that higher-income enrollees in Kaiser and Ross-Loos have more visits per year than lower-income enrollees, their own data show that this is also the case in the conventional plans. More importantly, low income PGP enrollees were much less likely to go without a physician visit, suggesting that the lower out-of-pocket cost in the PGPs more than compensates for any bureaucratic deterrent.

Taken together, these studies in large mature PGPs suggest some important differences in utilization patterns among regularly enrolled HMO members. While overall use of ambulatory services does not differ very much, especially when factors such as age, sex, and health status are held constant, the types of services vary. The middle class is much more likely to use the telephone to report symptoms. Walk-in visits and appointments, both of which are held during normal working hours, also appear easier for the middle class. While the unemployed may have a low time cost, the working poor probably have less flexible schedules than the middle class. Thus the poor are more likely to use the emergency room, and they are more likely to miss appointments without cancelling them.

The California Prepaid Health Plan Program Under Medicaid

The previous discussion provides substantial evidence that many HMOs have adequately served the poor and that their systems to constrain costs have not had noticeably adverse impacts. Yet, the history of HMOs and the poor contains one chapter with a different

tone—reports of fraud, abuse, poor quality, and inaccessible services by some prepaid health plans serving Medicaid beneficiaries in California. To gain a better understanding of how HMO-type services can lead to undesirable outcomes, it is useful to examine the California experience.

In 1970 California began a policy of encouraging the development of HMOs to serve its Medicaid population. These organizations, called prepaid health plans (PHPs), were supposed to save the state money be reducing unnecessary utilization and administrative costs of claims reimbursement. The then governor Ronald Reagan also attempted to control utilization of fee-for-service providers by limiting office visits and requiring prior authorization for hospital admissions. (See, e.g., Goldberg, 1975; California Department of Health, 1975; Chavkin and Treseder, 1977.) This strategy was also designed to make HMO enrollment more attractive to the Medicaid population, because no such external restrictions on utilization would be imposed on HMO enrollees. Implicit constraints were imposed because the HMO capitation rates were pegged to the average fee-for-service costs. Most observers agree that the initial program was poorly designed with problems arising from marketing abuses, excessive administrative costs and profits, and poor quality of care (Goldberg, 1975; D'Onofrio and Mullen, 1975; Moore and Breslow, 1972; California State Legislative Analyst, 1975; U.S. Senate Committee on Government Operations, Permanent Subcommittee on Investigations, 1975, 1978).

The principal charges concerning marketing stemmed from PHP use of door-to-door salespeople to enroll Medicaid eligibles. These people often posed as physicians or Medicaid officials and rarely explained fully the options available to consumers. Promises often far exceeded reality. Salespeople attempted to enroll only healthy individuals and told people with ongoing problems to stay in the fee-for-service system. These marketing abuses were at least partially attributable to the intense competition for enrollees among plans, which, in turn, was fostered by the ground rules laid down by the state (U.S. General Accounting Office, 1974). A PHP would sign a contract allowing it to receive a fixed monthly capitation fee of no more than 90 percent of the average fee-for-service cost of similar enrollees. No funds were available for start up or fixed costs, so it was imperative that the PHP enroll members as quickly as possible. Furthermore, in an effort to encourage competition and promote choice for Medicaid eligibles, the state took a laissez-faire attitude and authorized many plans to operate in the same service areas. While creating a situation in which there was a rush to enroll people, the state refused to release the names of Medicaid eligibles, so the PHPs had to use door-to-door salespeople paid on commission. There was also little control over PHP attempts to enroll only the healthy and to keep dissatisfied

members from disenrolling. The situation was almost perfectly designed to lead to an aggressive attempt to enroll healthy members.

There was almost no concern for even minimal quality audits in the early years of the PHP program, and only after June 1973 were medical professionals assigned by the state to audit PHPs (California State Legislative Analyst, 1975). A detailed investigation in 1974 of five PHPs in operation prior to December 1972 showed substantial variability in quality, with some significantly better than fee-for-service Medi-Cal (the California Medicaid program) and some significantly worse. Plans with a high proportion of non-Medi-Cal enrollees scored consistently better (Louis and McCord, 1974).

These differences in quality which appear related to the proportion of non-Medi-Cal enrollees are consistent with the notion that plans tend to provide comparable services for all enrollees and that the competitive environment has a major influence on the overall level of service. Large, mature HMOs have had to compete with conventional insurers and providers for the employed population. In contrast, some of the PHPs were established to serve only Medi-Cal prepaid enrollees while others were large fee-for-service Medi-Cal practices prior to converting to prepayment. In many areas, the poor have few alternatives—in 1972 less than 20 percent of Medi-Cal eligibles in Orange County saw "mainstream" providers (Auger and Goldberg, 1974).

This review suggests that in the past competition for employed populations has led to a high quality of care and service by HMOs. (These mature HMOs were also established by individuals committed primarily to the development of innovations in medical care rather than profitability, an issue to be discussed below.) But, competition per se is not necessarily good; the intense competition among PHPs for Medi-Cal enrollees encouraged their marketing abuses. Thus, future policy strategies should examine the nature and level of competition most likely to lead to socially desirable results.

Potential Effects Under Increased Competition

It is important to remember that the preceding discussion of HMO incentives and performance relates to a specific market environment and historical context. Most mature HMOs (i.e., those developed before the 1970s, which account for the bulk of total enrollment and most of the research studies) were started in an environment actively hostile to prepayment. The founders of these plans were often visionaries with strong beliefs concerning the superiority of prepayment and group practice (MacColl, 1966; Rothenberg et al., 1949; Wil-

liams, 1971). They often enlisted community support in their struggles with organized medicine and sought to provide a better service at reasonable cost. Few of these HMOs were operated on a for-profit basis and those aspects that were technically for profit, e.g., the physician groups, rarely seemed to act as if profit were a major goal. The market environment has also been relatively noncompetitive. Often only one HMO existed in an area. Although HMOs competed with conventional insurers for enrollees, tax subsidies for health insurance benefits and employer ignorance or unconcern made the market orderly and quiet. Current proposals to increase competition in the medical care market suggest that it may be inappropriate to extrapolate from the past.

Changes in the Competitive Environment

Several changes are proposed or in process to substantially increase competition. Various bills have been introduced to alter the tax laws to increase employee sensitivity to the cost of health insurance (Iglehart, 1981). Current law allows all employer contributions to be tax deductible for the employer and tax free to the employee. Many employers contribute more for more expensive plans. The proposed changes would place a cap on the employer contribution, require equal contributions, or give tax-free rebates to people choosing less expensive plans. Some employers provide "cafeteria style" benefit packages whereby employees receive a lump sum in fringe benefits to allocate among various health, life insurance, pension, vacation, and other options. The proposed changes in the tax law would increase the desirability of such cafeteria plans.

Employers have also become more aware of the cost of health insurance contributions and that this fringe benefit is a largely uncontrollable cost. Some employers have begun to alter their health insurance packages to increase cost control. One very popular approach is self-insurance and administrative-services-only arrangements in which the employer bears the risk and the carrier merely handles claims processing. The employer can then develop a substantial data base on cost and utilization and can exercise substantial control over reimbursement levels and payment flexibility.

New delivery systems and insurance arrangements are also beginning to enter the market. The older generation of PGPs and IPAs has been joined by new HMOs with a frankly entrepreneurial orientation. Some are local organizations, often controlled by a few physicians, while others are owned and operated by large statewide or national insurers. In some instances it has been alleged that these new plans have a very short (2–3 year) time horizon on profitability, rather than the long-run perspective focusing on service delivery attributed to the

older plans. (For example, the innovative United Healthcare plan sponsored by Safeco has been closed [Moore et al., 1980].) The currently high interest rates make imperative the rapid payback of start-up capital. New firms have entered the administrative services market with an emphasis on fast, inexpensive processing and a get-tough attitude toward providers. Such firms will sometimes offer a fixed fee, perhaps well below the physician's usual charge, and if that is not accepted, threaten interminable negotiations. Other plans are establishing "preferred provider organizations" in which selected cost-conscious providers are identified. If the enrollee goes to one of these providers, charges are covered in full, while for other providers standard copayments are required.

Simultaneously, the provider market is becoming increasingly competitive. Falling hospitalization rates are leading to increasing bed surpluses so hospitals are more willing to grant concessions (Kralewski and Countryman, 1982). For-profit hospital chains are expanding and sometimes concentrate on handling only certain types of cases. In some instances HMOs are owned by these for-profit chains. The supply of physicians is also increasing rapidly relative to the population and in some parts of the nation it is becoming difficult for physicians to start an independent practice and meet income expectations and debts incurred during training. This is leading to an increased willingness and sometimes eagerness by new, young physicians to accept a position with one of the new HMOs or other plans.

Possible Effects of Increased Competition

As competition increases, consumers and employers become more cost-conscious, and providers become more squeezed, various tactics may be used to ration resources inside and out of HMOs. Plans will probably attempt to control costs much more tightly than is currently the case. This may lead to the firing or dropping from provider lists of those physicians who appear to overutilize services (Physicians Health Plan, 1980). Both approaches have been used in selected plans. The question, of course, is whether the cuts are aimed only at the true "abuser" such as the physician who gives a vitamin B–12 shot to every patient. In a cost-cutting environment a plan may fail to distinguish the abusing physicians from those who have a sicker patient load. In either case, the patients of those physicians will lose a provider and the ones who will be most affected are the frequent users with chronic high-cost problems. Similar approaches are not difficult to imagine. For instance, plans generally cover all but "experimental" services. By not updating the coverage list frequently, patients desiring such new and usually expensive procedures will be encouraged to switch insurers (Rybin, 1981). Alternatively, the indications for a

service or its delivery method may be structured to shift certain users out of the plan. Many HMOs provide fairly comprehensive coverage of outpatient mental health care—often 20 or more visits per year, usually in a group therapy setting. While this may be sufficient for most acute problems, it may not be very attractive to patients who wish long-term insight-oriented individual therapy. Such people will switch to conventional insurers with outpatient mental health benefits. Plans could attract well-trained board-certified physicians yet make sure that, while the primary-care physicians are well liked by the enrollees, most of the specialists are foreigners with limited command of conversational English. Such a strategy would, if associated with low premiums, attract the young and healthy, and repel the old and sick.

Even more subtle tactics could be used by IPAs or insurers through manipulation of reimbursements. The allowable fees for certain types of expensive specialists could be held down, either as a matter of policy or by updating usual, customary, and reasonable fee screens less frequently for low-volume specialties (Showstack et al., 1979). Such an approach would increasingly shift costs to the patient and might even lead to a change of plan. A crucial factor in such approaches is that a small number of enrollees account for a large fraction of the costs (Roos and Shapiro, 1981; Eggers, 1981; McCall and Wai, 1981). Small changes that may affect only high users can have a major cost impact with little apparent effect on enrollment. Moreover, changes in administrative procedures or providers affecting this small segment are difficult or impossible to monitor from the outside.

Competitive strategies are even more likely to occur in marketing. Some of the procompetitive bills would eliminate minimum basic benefit packages. If such minimums are dropped, HMOs and insurers could compete for low-cost enrollees by offering benefits attractive to the healthy, like eyeglasses, while limiting coverage for services required by high-cost enrollees. Plans are likely to target their enrollment toward the healthier groups, and within groups toward healthier and less expensive individuals.

Employers may also structure their benefit packages to reduce their own costs. For instance, it is reported that one Washington, D.C.-based firm has designed its health-benefits option so employees married to federal workers will choose coverage under their spouse's plan, thus reducing costs to the firm. (While such anecdotal evidence is difficult to generalize, it is indicative of the types of behavior already observed which are likely to become more common. The examples are drawn from interviews during which anonymity was assured.) In another situation, a firm with a self-insured plan and a community-rated HMO option had personnel staff counsel people with health problems to join the HMO—and thus lower the firm's costs. Some employers with

administrative-services-only contracts have their carriers identify high-utilizing employees for special counseling. It is only a small additional step to pressure such employees to quit. Preemployment physical exams can be used to screen out not only people unable to perform their jobs, but those who may have costly medical expenses.

These are just a few of the possible ways by which increased competition may lead to a situation in which the young and healthy receive good low-cost coverage and the old and sick face expensive inadequate coverage—just the opposite of the risk-pooling implicit in socially desirable insurance. Many more tactics are likely to be developed as competitive pressures increase. While the old and sick are more likely to be poor, other factors will also increase the burden of such a competitive strategy for poor and female workers. Unions are likely to resist the fragmentation envisioned in this scenario and employers of upper-income workers are also less likely to adopt penny-pinching strategies. For instance, one electronics firm was said to avoid aggressive intermediaries because it did not want to offend its engineers with a hard line on claims. Low-income and non-union workers have much smaller employer contributions and are likely to be especially sensitive to the self-selection inducements of a competitive strategy (Taylor and Lawson, 1981).

If Medicare and Medicaid are brought into the competitive strategy similar outcomes are possible. The experience with the "Medigap" policies suggests that many of the elderly are not informed consumers of health insurance; the poor are apt to be even more vulnerable. Depending upon the value of the "voucher" offered by the government, the Medicare and Medicaid beneficiaries could be very attractive markets. On average, their medical expenditures (and implicit premiums) are high, but again, a small fraction accounts for a very large fraction of the expenditures (Roos and Shapiro, 1981). Furthermore, in contrast to employed groups, high expenditures among Medicare and Medicaid beneficiaries are probably more likely to be associated with chronic conditions than sudden trauma so the importance of preexisting physician ties is even stronger. It may also be easier to attract people with certain chronic, but low-cost conditions while discouraging others. For instance, a plan may emphasize a program for the blind, who are relatively low-cost patients, while making access difficult to cardiologists. The relative absence of informed agents, such as union negotiators, may make such strategies particularly easy if there is no government monitoring of the process.

Ethical and Policy Implications of Rationing of Medical Care in HMOs

The previous sections have pointed out that the systems of controlling utilization or rationing in HMOs are rather different than they might

appear at first glance. While fees are not used to influence consumer demand, the burden is placed only partially on time prices. The seemingly endless queues reported in England are not evident in most HMOs. Instead, there is a different tradeoff between waiting time and appointment delay in PGPs and, while physical accessibility is more difficult in some ways, it is easier in others. Rationing tends to occur in much more subtle ways. Controls over services provided take place through an environment that encourages the medical staff to be more conservative in its use of medical resources. Hard and fast rules on when a patient should be admitted have not been uncovered. Instead, although clinicians make decisions on a case-by-case basis, the overall pattern of decisions is more conservative. This may be the result of changing the practice patterns of physicians or of attracting those physicians who already have conservative practice patterns. Perhaps a more important type of rationing occurs through the process by which consumers enroll in HMOs. A complex mixture of decisions by government, insurers, HMOs, employers, unions, and the individuals themselves results in HMO options being unavailable to some, expensive to others, and attractive to a relatively small proportion of the nation.

The effects of these rationing devices on various subgroups of the population are similarly complex. By and large, the poor and other disadvantaged groups enrolled in HMOs probably have better access to services than would be the case for them in the conventional system. Similarly, the differences in utilization among socioeconomic groups in HMOs are far smaller than in conventional settings. This is not to say that all differences are erased or that utilization of services is strictly in proportion to need. However, many of the remaining differences seem attributable to difficulties faced by the poor in all settings, rather than to specific hurdles imposed by the HMOs. Somewhat broader differential effects can be seen in those who have the option of joining HMOs. However, this rationing through the availability of enrollment options is to a substantial degree a reflection of market forces and government policies.

Ethical Issues in Approaches Used by HMOs to Constrain Utilization

Two major ethical issues arise when considering the methods used by HMOs to allocate medical services: 1) whether the economic incentives present in an HMO create unethical behavior by their physicians, and 2) whether the allocation system appropriately reflects medical need. Unethical physician behavior may arise from the loss of the physician as an impartial and trusted agent for the patient. This important issue lies at the heart of traditional opposition to the

corporate practice of medicine. Abstracting from the rhetoric of such discussions, one can say that the appropriate level of treatment within the current market environment is that level of treatment which the fully informed patient would choose after weighing its risks and benefits, as well as its financial costs (Pauly, 1979). (To simplify matters, we will assume the patient's insurance coverage is equivalent under both systems. Various redistributive schemes would add complexity to the discussion but not alter its major conclusion.) A physician acting as a perfect agent for this patient would reach a similar conclusion. A physician who is also a provider of care to this patient may no longer be a perfect agent and may be swayed by the economic incentives inherent in the payment scheme. If paid fee-for-service, he or she will have an incentive to offer more services, while if paid on a salary or capitation basis, there is an incentive to offer fewer services. The question is whether it is better to err on the high or low side of the correct amount. Too many services increase not only cost but risk.

Unfortunately, it is rarely the case that the patient's problem is diagnosed with such certainty and that a course of treatment is so clear that the "correct" amount is known. Instead, there is generally some uncertainty about the true nature of the disease and often there are several courses of treatment with no clear objective evidence as to which is best for the particular patient (Wennberg et al., 1980). The patient does not have the expertise to really evaluate the potential choices, and I suspect that few physicians are able to deal explicitly with the uncertainty. In fact, even consciously recognizing the relative "softness" of the decision may be such a burden that many physicians develop a "standard approach" to dealing with particular problems. Some may choose a very aggressive testing and treatment approach while others may choose to "wait and see" with minimal intervention. Yet both approaches may yield similar outcomes, especially since our measures of outcome are rather insensitive. The similarity in results will reinforce each practitioner in his or her chosen approach. Wide variations in practice styles have been observed among independent fee-for-service practitioners, as well as within apparently similar physicians in both fee-for-service and prepaid group practices (Wennberg and Gittelsohn, 1973; Schroeder et al., 1973; Roos and Roos, 1981).

While there may be a wide distribution across physicians in how they might treat a particular problem, fee-for-service physicians are more likely to be concentrated at the high service end and prepaid physicians at the low service end. This may reflect the role of incentives in altering their clinical decision making. Alternatively, it may reflect a process of selection. Physicians who order fewer tests and procedures and hospitalize less frequently—perhaps because that is the way they were trained—will find it difficult to earn as much in a fee-for-service

environment as will more aggressive physicians. Similarly, physicians with an aggressive approach will find a prepaid environment less rewarding than fee-for-service (McClure, 1982). Yet each type may be practicing within the generally accepted range of clinical behavior.

There are very few studies of different practice patterns in fee-for-service and prepaid settings that examine what is or would be done for identical patients. Overall measures of the use of services are confounded by patient copayments, differences in case mix, and other factors. One very recent study queried board-certified cardiologists on how they would treat a series of patients described in brief case histories (Hlatky et al., 1981). As might be expected, the independent fee-for-service physicians on average were much more likely than the prepaid group-practice cardiologists to recommend invasive testing and cardiac bypass surgery. However, the recommendations of the HMO physicians were similar to those of university-based cardiologists. In general, recommendations were comparable across practice settings for patients with the most severe manifestations, and major discrepancies occurred for patients with only one or two blocked arteries, a situation in which there is no clear consensus on appropriate treatment in the research literature (McIntosh, 1981; Kouchoukos, 1981). Furthermore, some fee-for-service cardiologists were *more* conservative than the PGP cardiologists, supporting the notion that while the average recommendation may differ significantly, the PGP physicians are within a range acceptable to fee-for-service providers. In general, current studies of the quality of care in HMOs indicate standards at least as high as those in the fee-for-service community (Cunningham and Williamson, 1980; Luft, 1981).

The second ethical issue is whether the systems used by HMOs to ration their services result in an allocation that reflects medical need. HMOs seem to perform well in allocating care *among their enrollees* largely according to need, rather than ability to pay. In this respect, their performance is superior to the conventional system in which copayments often serve as greater deterrents for those with lower incomes.

From a broader perspective, however, one may ask whether *access to health plans* should be independent of ability to pay. There are two problems raised by this question. The first relates to the inability of the poor to pay for health insurance premiums. This problem can, in theory, be solved by giving the poor and low-income workers a voucher scaled to their income and family size that can be used to purchase basic health insurance. If they wish a more expensive plan they will have to pay the extra cost out-of-pocket, just as would be the case for the middle class. Everyone would be placed on an equal footing and access to good basic coverage would be available to all. Such an approach is outlined in Alain Enthoven's "Consumer Choice

Health Plan" (1980) and is incorporated in the "Project Health" system in Multnomah County, Oregon (Lewis, 1979; Multnomah County, Oregon, Project Health Division, 1977). The "Project Health" system provided not only a basic voucher, but provided subsidies so that low-income enrollees did not have to pay the full marginal cost of more expensive plans.

The second problem in assuring equal access to health plans arises from the higher cost of covering sick people and the resulting incentives for plans to avoid or push out such enrollees. As indicated earlier, increasing competition is likely to increase the attempts by plans—both HMOs and conventional insurers—to enroll only the healthy. This suggests that people with high needs for medical care will have to purchase coverage at substantial out-of-pocket cost or else they will be forced out of the system. The proposed scenario is analogous to the change in insurance during the 1950s and 1960s which led to the virtual replacement of community rating by experience rating. An area would start out with a uniform rate for all persons. Insurers would identify firms with younger-than-average employees and offer a lower rate. As low-cost firms dropped out of the pool, the premium for those remaining would increase and the process accelerate. The result was a wide variation in rates charged to different firms with some gaining and others losing. Increased competition in the 1980s is likely to result in conventional insurers and HMOs trying to attract low-cost *individuals* within firms, leaving the older or sicker persons to high-cost plans. Furthermore, as the California PHP example and other early indications suggest, the competitive techniques that may be used are not always aboveboard.

Implications for Health Policy

The crucial policy issue that derives from this discussion is how can a competitive system be designed to allow for diversity and encourage innovation and efficiency without getting caught by an analogue to Gresham's law in which cheap plans attract the healthy and exclude the sick. Unfortunately, it is far easier to raise concerns and identify problems than it is to propose solutions, but a few suggestions will be offered to promote discussion.

The basic approach in these suggestions is to reduce the benefits to those plans that select healthy enrollees and increase the penalties they would pay for pursuing such strategies. The first step would be to require a rather broad basic-benefit package for all plans so that people with specific health problems would not be deterred by benefit exclusions. Plans would be encouraged to compete primarily on the basis of price rather than coverage options. This will also make it easier for consumers to evaluate alternatives (Luft, 1982).

The major difficulty with any competitive approach is that the effective premiums for the sick must be above those for the healthy; otherwise plans will develop strategies to avoid enrolling people whose expected costs exceed the premium. If it is possible to identify with reasonable accuracy those people at risk for high expenses, they can be categorized and charged a higher premium. This would entail a system analogous to, but much more complicated than, the point system used by carriers to adjust insurance premiums for drivers with poor accident or traffic violation records. A social decision might be made that most health problems are not the individual's responsibility, so government vouchers or employer contributions would be adjusted to reflect the individual's risk category. While a rather complex system might result, the philosophical precedent is already implicit in the current willingness to pay high premiums for Medicare coverage of the elderly, persons with kidney failure, and the permanently disabled.

The actuarial feasibility of designing risk-adjusted premiums and vouchers is quite another problem. Making the premium a smooth function of several factors such as age, sex, family medical history, and the like might avoid many of the problems associated with the yes/no decisions on disability eligibility. Such a system would make the sick and healthy equally desirable to plans because each brings a premium commensurate with expected costs. Very careful screening by plans might still produce some selection, but most of the benefit to plans of such approaches would be lost. The vouchers could be set at some constant fraction of the risk-adjusted premiums if price sensitivity were desired. Furthermore, the plan's actual premium might be above or below the voucher level depending on the plan's efficiency.

Enthoven (1980) includes a more restricted version of this notion in his discussion of actuarial risk categories. My major difference with him is in emphasis. I think we agree that without such adjustments competition will result primarily in a rush for healthy enrollees. However, I am much more skeptical than Enthoven about the technical feasibility of devising a sufficiently sensitive categorizing scheme, but this is a research question which has not to my knowledge even been explored in a preliminary fashion. The government's willingness to maintain its contribution levels for the most costly categories is a crucial political question that must also be addressed. This becomes particularly important if high-cost patients tend to enroll in certain plans (e.g., those utilizing tertiary-care hospitals) and the plan's high costs are attributed to inefficiency rather than to patients with more complicated problems.

The rewards a plan can reap by pushing out its high-cost enrollees can also be reduced. It may be possible to develop monitoring systems whereby people who change plans report on the problems they experienced and whether they felt pushed out. For instance, the Cali-

fornia Public Employees Retirement System, which provides multiple-choice options for state employees, sends a questionnaire to all enrollees switching plans. Because such a system must be sensitive enough to identify important but infrequent complaints, it may be necessary to pool the data from many employee groups to obtain a sufficient sample. Such a system might be analogous to that used to evaluate automobile failures nationwide.

The incentives to keep nonusers while disenrolling users might be blunted by allowing enrollees to build up credits over several years for below-average use. If they switch to another plan, say after incurring an illness, some of those credits would be transferred and drawn down by the new plan. If the previous plan induced them to stay, it could use some of the credits to offset the patient's higher expenses. Although such an approach may be criticized as contrary to general insurance principles, it is designed to counter strong market incentives that may destroy the basic foundation of risk-spreading insurance.

It should be clear that these are only preliminary ideas presented to encourage discussion. They rest upon the belief that HMOs currently provide an important alternative to the conventional system, but that potential changes in the market environment may result in rather objectionable behavior by HMOs and other health plans. Whether it will be possible to design the appropriate policies to encourage desirable, and discourage undesirable, performance rests on improving our knowledge of how the medical care system works and on creating the correct incentives.

References

Arrow, K.J. 1963. Uncertainty and the Welfare Economics of Medical Care. *American Economic Review* 53(5):941–973.

Auger, R., and Goldberg, V. 1974. Prepaid Health Plans and Moral Hazard. *Public Policy* 22:353–393.

Berki, S.E., and Ashcraft, M.L.F. 1980. HMO Enrollment: Who Joins and Why: A Review of the Literature. *Milbank Memorial Fund Quarterly/Health and Society* 58(4):588–632.

California Department of Health. 1975. Prepaid Health Plans: The California Experience. In U.S. Senate, Committee on Government Operations, Permanent Subcommittee on Investigations, *Prepaid Health Plans,* Hearings, March 13–14, 1975, 94th Congress, 1st Session. Washington, D.C.: U.S. Government Printing Office.

California State Legislative Analyst. 1975. A Review of the Regulation of Prepaid Health Plans by the California Department of Health, Sacramento, November 15, 1973. In U.S. Senate, Committee on Government Operations, Permanent Subcommittee on Inves-

tigations, *Prepaid Health Plans*, March 13–14, 1975, 94th Congress, 1st Session. Washington, D.C.: U.S. Government Printing Office.

Chavkin, D.F., and Treseder, A. 1977. California's Prepaid Health Plan Program: Can the Patient Be Saved? *Hastings Law Journal* 28 (January):685–760.

Collen, M.F., Dales, L.G., Friedman, G.D., Flagle, C.D., Feldman, R., and Siegelaub, A.B. 1973. Multiphasic Checkup Evaluation Study. 4. Preliminary Cost Benefit Analysis for Middle-Aged Men. *Preventive Medicine* 2(2):236–246.

Coltin, K., Neisuler, R., and Lurie, R.S. 1978. Utilization of Preventive Services by Poor, Near-Poor, and Non-Poor Members of an HMO. Presented at 106th Annual Meeting of the American Public Health Association, October 16. (Unpublished.)

Cook, W.H. 1971. Profile of the Permanente Physician. In Somers, A.R., ed., *The Kaiser-Permanente Medical Care Program: A Symposium*. New York: Commonwealth Fund.

Cunningham, F.C., and Williamson, J.W. 1980. How Does the Quality of Health Care in HMOs Compare to That in Other Settings—An Analytic Literature Review: 1958 to 1979. *Group Health Journal* 1(1):4–25.

D'Onofrio, C.N., and Mullen, P.D. 1977. Consumer Problems with Prepaid Health Plans in California. *Public Health Reports* 92(2):121–134.

Eggers, P. 1981. Pre-Enrollment Reimbursement Patterns of Medicare Beneficiaries Enrolled in "At Risk" HMOs. Health Care Financing Administration, Office of Research Demonstration and Statistics. (Unpublished draft.)

Enthoven, A. 1980. *Health Plan: The Only Practical Solution to the Soaring Cost of Medical Care*. Reading, Mass.: Addison-Wesley.

Freeborn, D.K., Pope, C.R., Davis, M.A., and Mullooly, J.P. 1977. Health Status, Socioeconomic Status, and Utilization of Outpatient Services for Members of a Prepaid Group Practice. *Medical Care* 15(2):115–128.

———, Greenlick, M.R., Mullooly, J.P., and Burnham, V. 1980. *The Kaiser-Permanente Neighborhood Health Center Project: Ambulatory Care Utilization Patterns*. Health Services Research Center, Research Report Series, Report No. 5. Portland, Oreg.: Kaiser-Permanente Medical Care Program.

Fuller, N., and Patera, M. 1976. *Report on a Study of Medicaid Utilization of Services in a Prepaid Group Practice Health Plan*. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Public Health Service, Bureau of Medical Services.

Gaus, C., Cooper, B.S., and Hirschman, C.G. 1976. Contrasts in HMO and Fee-For-Service Performance. *Social Security Bulletin* 39(5):3–14.

Gertman, P.M. 1974. Physicians as Guiders of Health Services Use. In Mushkin, S.J., ed., *Consumer Incentives for Health Care*. New York: Prodist.

Goldberg, V.P. 1975. Some Emerging Problems of Prepaid Health Plans in the Medi-Cal System. *Policy Analysis* 1(1):55–68.

Group Health Cooperative of Puget Sound. 1970. Non-Appearance of Appointed Enrollees for Care among Public Assistance Enrollees for Comprehensive Prepaid Medical Care. *GHC Research Note* No. 9, summarized in Vignola and Strumpf, 1980.

Held, P.J., and Reinhardt, U. 1980. Prepaid Medical Practice: A Summary of Findings from a Recent Survey of Group Practices in the United States—A Comparison of Fee-for-Service and Prepaid Groups. *Group Health Journal* 1(Summer):415.

Hetherington, R.W., Hopkins, C.E., and Roemer, M.I. 1975. *Health Insurance Plans: Promise and Performance.* New York: Wiley-Interscience.

Hirschman, A.O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations and States.* Cambridge, Mass.: Harvard University Press.

Hlatky, M., Botvinick, E., and Brundage, B. 1981. A Controlled Comparison of Cardiac Diagnostic Test Use in a Health Maintenance Organization. Presented at Annual Meeting of Robert Wood Johnson Clinical Scholars, San Antonio, Texas, November 11–14. (Unpublished.)

Hurtado, A.V., Greenlick, M.R., and Colombo, T.J. 1973. Determinants of Medical Care Utilization: Failure to Keep Appointments. *Medical Care* 10(3):189–198.

Iglehart, J.K. 1981. Drawing the Lines for the Debate on Competition. *New England Journal of Medicine* 305(5):291–296.

Johnson, R.E., and Azevedo, D.J. 1979. Comparing the Medical Utilization and Expenditures of Low Income Health Plan Enrollees with Medicaid Recipients and with Low Income Enrollees Having Medicaid Eligibility. *Medical Care* 17(9):953–966.

Kouchoukos, N.T. 1982. *An Overview Paper—a Surgeon's Perspective on Coronary Artery Bypass Surgery.* Washington, D.C.: National Center for Health Care Technology.

Kralewski, J.E., and Countryman, D.D. The Interaction of Community Hospitals with HMOs: A Case Study of the Minneapolis/St. Paul Metropolitan Area. *Health Care Financing Review,* in press.

Lewis, R. 1979. Oregon County Buying Prepaid Care for the Needy. *American Medical News* (July 17):1,14,15.

Louis, D.Z., and McCord, J.J. 1974. *Evaluation of California Prepaid Health Plans: Executive Summary.* Santa Barbara: General Research Corp.

Luft, H.S. 1978. How Do Health Maintenance Organizations Achieve Their "Savings"?: Rhetoric and Evidence. *New England Journal of Medicine* 298 (June 15):1336–1343.

———. 1981. *Health Maintenance Organizations: Dimensions of Performance.* New York: Wiley-Interscience.

———. 1982. On the Potential Failure of Good Ideas: An Interview with the Originator of "Murphy's Law." *Journal of Health Politics, Policy and Law,* in press.

———, Feder, J., Holahan, J., and Lennox, D.K. 1980. Health

Maintenance Organizations. In Feder, J., Holahan, J., and Marmor, T., eds., *National Health Insurance: Conflicting Goals and Policy Choices*. Washington, D.C.: Urban Institute.

MacColl, W.A. 1966. *Group Practice and Prepayment of Medical Care.* Washington, D.C.: Public Affairs Press.

McCall, N., and Wai, H.S. 1981. *An Analysis of the Use of Medicare Services by the Continuously Enrolled Aged.* Menlo Park, Calif.: SRI International.

McClure, W. 1982. Toward Development and Application of a Qualitative Theory of Hospital Utilization. *Inquiry,* in press.

Matlock, M.A. 1980. Birth Rate Begets HMO Cancellation. *Business Insurance* (December 22):1,31.

McIntosh, H.D. 1981. *Overview of Aortocoronary Bypass Grafting for the Treatment of Coronary Artery Disease: An Internist's Perspective.* Washington, D.C.: National Center for Health Care Technology.

Mechanic, D. 1976. *The Growth of Bureaucratic Medicine.* New York: Wiley-Interscience.

Meier, G.B., and Tillotson, J. 1978. *Physician Reimbursement and Hospital Use in HMOs.* Excelsior, Minn.: InterStudy.

Moore, S.H., Martin, D.P., Richardson, W.C., and Reidel, D.C. 1980. Cost Containment through Risk-Sharing by Primary Care Physicians: A History of the Development of United Healthcare. *Health Care Financing Review* 1(4):1–14.

Moore, T.G., Jr., and Breslow, L. 1972. California Medical Group Evaluation Report. California Council for Health Plan Alternatives, December 6. In U.S. Senate, Committee on Government Operations, Permanent Subcommittee on Investigations, *Prepaid Health Plans,* Hearings, March 12–14, 1975, 94th Congress, 1st Session. Washington, D.C.: U.S. Government Printing Office.

Multnomah County Oregon, Project Health Division. 1977. *National Health Insurance through State/Regional/Local Government and Private Sector Health Care Brokerage.* Portland, Oreg.: Department of Human Services.

Newhouse, J.P., Manning, W.G., Morris, C.N., Orr, L.L., Duan, N., Keeler, E.B., Leibowitz, A., Marquis, K.H., Marquis, M.S., Phelps, C.E., and Brook, R.H. 1981. Some Interim Results from a Controlled Trial of Cost Sharing in Health Insurance. *New England Journal of Medicine* 305(25):1501–1507.

Nolan, R.L., Schwartz, J.L., and Simonian, K. 1967. Social Class Differences in Utilization of Pediatric Services in a Prepaid Direct Service Medical Care Program. *American Journal of Public Health* 57(1):34–47.

Pauly, M.V. 1979. What Is Unnecessary Surgery? *Milbank Memorial Fund Quarterly/Health and Society* 57(1):95–117.

Phelan, J., Erickson, R., and Fleming, S. 1970. Group Practice Prepayment: An Approach to Delivering Organized Health Services. Part II. *Law and Contemporary Problems* 35(4):796–816.

Physicians Health Plan. 1980. *The Physicians Health Plan of Minnesota: A Case Study of Utilization Controls in an IPA.* DHHS Publication

No. (PHS) 80-50128. Washington, D.C.: U.S. Government Printing Office.

Pope, C.R., Yoshioka, S., and Greenlick, M.R. 1971. Determinants of Medical Care Utilization: The Use of Telephone for Reporting Symptoms. *Journal of Health and Social Behavior* 12 (June):155–162.

Richardson, W.C., Boscha, M.V., Diehr, P.K., Drucker, W.L., Efird, R.A., Green, K.E., LoGerfo, J.P., McCaffree, K.M., Richardson, W.C., Shortell, S.M., Weaver, B.L., and Williams, S.J. 1977. *The Seattle Prepaid Health Care Project: Comparison of Health Services Delivery*. PB 267488–SET. Springfield, Va.: National Technical Information Service.

Roos, N.P., and Roos, L.L. 1981. High and Low Surgical Rates: Risk Study Factors for Area Residents. *American Journal of Public Health* 71(6):591–600.

Roos, N.P., and Shapiro, E. 1981. The Manitoba Longitudinal Study on Aging: Preliminary Findings on Health Care Utilization by the Elderly. *Medical Care* 19(6):644–657.

Rothenberg, R.E., Pickard, K., and Rothenberg, J.E. 1949. *Group Medicine and Health Insurance in Action*. New York: Crown.

Rybin, V. 1981. Cry for Help: Can HMOs Afford to Listen to Their Patients? *St. Paul Sunday Pioneer Press* (November 15).

Salkever, D.S., German, P.S., Shapiro, S., Horky, R., and Skinner, E.A. 1976. Episodes of Illness and Access to Care in the Inner City: A Comparison of HMO and non-HMO Populations. *Health Services Research* (Fall):252–270.

Schroeder, S.A., Kenders, K., Coopers, J.K., and Piemme, T.E. 1973. Use of Laboratory Tests and Pharmaceuticals: Variation among Physicians and Effects of Cost Audit on Subsequent Use. *Journal of the American Medical Association* 225(8):969–973.

Scitovsky, A.A., Benham, L., and McCall, N. 1979. Use of Physician Services under Two Prepaid Plans. *Medical Care* 17(5):441–460.

Showstack, J.A., Blumberg, B.D., Schwartz, J., and Schroeder, S.A. 1979. Fee-for-Service Physician Payment: Analysis of Current Methods and Their Development. *Inquiry* 16(3):230–246.

Shragg, H., Fagenbaum, M.E., Kovner, J.W., Caro, H.M., and Bunting, E.D. 1973. Low Income Families in a Large Scale Prepaid Group Practice. *Inquiry* 10(2):52–60.

Taylor, A.K., and Lawson, W.R., Jr. 1981. Employer and Employee Expenditures for Private Health Insurance. *NCHSR National Health Care Expenditures Study Data Preview 7*. DHHS Publication No. (PHS) 81-3297. Washington, D.C.: U.S. Government Printing Office.

Thompson, R.S. 1979. Approaches to Prevention in an HMO Setting. *Journal of Family Practice* 9(1):71–82.

U.S. General Accounting Office. 1974. *Better Controls Needed for Health Maintenance Organizations under Medicaid in California*. Washington, D.C.: General Accounting Office. September 10.

U.S. Senate. Committee on Governmental Affairs, Permanent Subcommittee on Investigations. 1978. *Prepaid Health Plans and Health Maintenance Organizations*, 95th Congress, 2nd Session,

Report 95-749. Washington, D.C.: U.S. Government Printing Office.

U.S. Senate. Committee on Government Operations, Permanent Subcommittee on Investigations. 1975. *Prepaid Health Plans,* Hearings, March 13–14, 1975, 94th Congress, 1st Session. Washington, D.C.: U.S. Government Printing Office.

Vignola, M.L., and Strumpf, G.B. 1980. *Medicaid Beneficiaries in Health Maintenance Organizations: Utilization, Cost, Quality, Legal Requirements—An Annotated Bibliography.* DHHS Publication No. 10012. Washington, D.C.: American Public Welfare Association and U.S. Health Care Financing Administration.

Weiss, J.E., and Greenlick, M.R. 1970. Determinants of Medical Care Utilization: The Effects of Social Class and Distance on Contacts with the Medical Care System. *Medical Care* 8(6):456–462.

Wennberg, J.E., and Gittelsohn, A. 1973. Small Area Variations in Health Care Delivery. *Science* 182(4117):1102–1108.

Wennberg, J.E., Bunker, J.P., and Barnes, B. 1980. The Need for Assessing the Outcome of Common Medical Practices. In Breslow, L., Fielding, J.E., and Lave, L.B., *Annual Review of Public Health,* Vol. I. Palo Alto: Annual Reviews, Inc.

Williams, G. 1971. *Kaiser-Permanente Health Plan: Why It Works.* Oakland, Ca.: Henry J. Kaiser Foundation.

Legal Implications of Allocation Policies

V

Rationing Medical Resources: A Constitutional, Legal, and Policy Analysis

James F. Blumstein*

I. Introduction

Until the 1970s, the major health care issues in the national political arena were improved quality and increased access to medical care.[1] Governmental support for biomedical research grew munificent, and technological progress improved the quality of medical care in many areas.[2] Programs such as Medicare[3] for the elderly and Medicaid[4] for the poor provided governmental financial backing for the delivery of medical services to defined groups of needy and worthy persons. These programs succeeded in expanding access to the medical care system, considerably narrowing, or in some cases even eliminating, the perceived "access gap."[5]

Progress in improving the quality and availability of medical care, however, has in recent years brought another matter to the forefront of health policy debates—the overall costs of health care.[6] In the last decade, the public has riveted its attention on the problem of cost contain-

* Professor of Law, Vanderbilt Law School; Senior Research Associate, Vanderbilt Institute for Public Policy Studies. B.A. 1966, M.A. 1970, LL.B. 1970, Yale University.

An earlier version of this Article was prepared for the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. The views represent those of the author, not those of the Commission.

1. *See* T. MARMOR & J. MARMOR, THE POLITICS OF MEDICARE (1973); Marmor, Feder & Holahan, *Introduction*, in NATIONAL HEALTH INSURANCE: CONFLICTING GOALS AND POLICY CHOICES 1, 7-12 (J. Feder, J. Holahan & T. Marmor eds. 1980).

2. *See* Wagner & Zubkoff, *Medical Technology and Hospital Costs*, in HOSPITAL COST CONTAINMENT 263 (M. Zubkoff, I. Raskin & R. Hanft eds. 1978).

3. Social Security Amendments of 1965, Pub. L. No. 89-97, § 1801, 79 Stat. 286, 291 (codified at 42 U.S.C. § 1395 (1976 & Supp. III 1979)).

4. *Id.* § 1901, 79 Stat. 286, 343 (codified at 42 U.S.C. § 1396 (1976 & Supp. III 1979)).

5. *See* L. ADAY, R. ANDERSEN & G. FLEMING, HEALTH CARE IN THE U.S.: EQUITABLE FOR WHOM? (1980); K. DAVIS & C. SCHOEN, HEALTH AND THE WAR ON POVERTY: A TEN YEAR APPRAISAL 41 (1978); F. SLOAN & J. BENTKOVER, ACCESS TO AMBULATORY CARE AND THE U.S. ECONOMY 147 (1979); Davis & Reynolds, *The Impact of Medicare and Medicaid on Access to Medical Care*, in THE ROLE OF HEALTH INSURANCE IN THE HEALTH SERVICES SECTOR 391, 392 (R. Rosett ed. 1976).

6. *See* Gibson & Waldo, *National Health Expenditures, 1980*, 3 HEALTH CARE FINANCING REV. 1 (1981). *See also* Marmor, Feder & Holahan, *supra* note 1, at 11.

Published in 59 TEXAS L. REV. 1345 (1981). Copyright © 1981 by the Texas Law Review. Reprinted by permission.

ment in health care and what to do about it. Discussions of health policy issues now commonly begin by reciting cost escalation data,[7] including such items as per capita expenditure levels and, most significantly, the percentage of the gross national product devoted to health care.[8] The rationing of care—by both the government and private decisionmakers—is sometimes viewed as one possible method of combatting the problem of cost escalation.[9]

The central question addressed in this Article is the effect of legal constraints on the process of rationing medical care and the substantive criteria for such rationing. A full consideration of the legal issues associated with rationing requires an analysis of the reasons that cost containment and resource allocation have become important issues in health policy. This Article examines the policy implications of the development of criteria for allocating scarce medical resources and then focuses on constitutional concerns, both procedural and substantive. It concludes by examining the relationship between rationing and medical malpractice law.

II. Defining Problems and Identifying Issues

To acknowledge that cost containment in medical care is a "problem" is much more significant than is often fully appreciated. While we are all too accustomed to complaining about the general level of inflation, we rarely concern ourselves with increases in aggregate levels of expenditure in any single industry. It would be extraordinarily rare to debate, as a public policy issue, the question of whether we devote too much of our national wealth to television sets as compared to some other productive use. Such debates, however, are common in the health sector.

The difficulty in answering this type of question stems from our lack of technical expertise and, more importantly, from the lack of any consensus on ultimate goals and allocative criteria. The heavy overlay of ideology and symbolism in the health sector exacerbates the general difficulty of the task.[10] A balanced discussion of the legal issues associated with rationing medical care must begin with an examination of why this atypical public policy issue has arisen in the context of health care.

A. Resource Allocation Decisions in a Market System

In a market economy, questions concerning appropriate resource allocation in the aggregate are not often posed in public policy terms. The impersonal market allocates resources in accordance with the consumer preferences expressed in numerous low-level exchange transac-

7. *See, e.g.*, A. ENTHOVEN, HEALTH PLAN xv-xvii, 13-32 (1980).

8. In 1980, health expenditures amounted to $1,067 per person. About 42.2% of that amount ($450) came from public funds. Total health care expenditures for 1980 are estimated at $247 billion, approximately 9.4% of the gross national product. *See* Gibson & Waldo, *supra* note 6.

9. *See, e.g.*, Schwartz, *We Need to Ration Medicine*, NEWSWEEK, Feb. 8, 1982, at 13.

10. *See* Zubkoff & Blumstein, *The Medical Marketplace: Health Policy Formulation in Consideration of Economic Structure*, in 2 NATIONAL COMMISSION ON THE COST OF MEDICAL CARE, 1976-1977, at 73, 79-81 (1978).

tions.[11] With respect to private goods and services, resource allocation in the aggregate is, therefore, not usually a public policy agenda item in a smoothly functioning market system. Resource scarcity exists, of course, in any economic system, and resources optimally should be allocated to their most efficient use. A salient characteristic of a market economy, however, is that such macroeconomic allocation issues need not be addressed in these global terms. Since decisions are decentralized to millions of individual decisionmaking units, each presumably efficient and with its own set of preferences and constraints, aggregate resource allocation decisions do not derive from a conscious, centralized process of determining needs and setting priorities at the front-end. Rather, aggregate (macro) allocation decisions are only inferred from adding up the consequences of each individual, decentralized (micro) transaction.

Under traditional notions of a functioning market system, then, the "problem" of cost containment and the "need" for "rationing" scarce resources do not arise as single aggregative decisions. The market generates prices that, in theory, reflect and communicate real resource scarcity. The decentralized decisionmaking units, both firms and households, are then able to adjust their behavior in light of the price signals generated by the system. Only after national income accountants have subsequently aggregated these private, atomistic decisions can we determine how society in fact chose to allocate its collective scarce resources in any given accounting period. In the market economy, therefore, data on resource allocation are derived after the fact by statistical inference. The collective *problem* of resource allocation is thus foreign to a properly functioning private market, because resource allocation decisions are *observed* in a statistical sense, not *made* in a consciously collective sense.

B. The Role for Government in a Market System

The market system is best characterized as a framework in which private decisionmakers act in accordance with the cues provided by the price system. In order for the system to function effectively, the participants must have access to information upon which to make rational decisions.[12] Artificial barriers to consumer choice impede the efficient functioning of the market process. Most importantly, a basic assumption of the market system is that the parties to any given exchange transaction bear the costs of and secure the benefits from that transaction. Other people, and society generally, are presumed to have little or no interest in each of these atomistic decisions.[13]

The conditions necessary for the smooth functioning of the market do not always obtain. Even the most ardent proponents of the market system recognize that a role for government in the economy is often

11. *See generally* J. HIRSHLEIFER, PRICE THEORY AND APPLICATIONS (1980). For a discussion from a health care perspective, see P. FELDSTEIN, HEALTH CARE ECONOMICS 405-22 (1979). *See also* Zubkoff & Blumstein, *supra* note 10, at 92-93.

12. *See generally* J. NEWHOUSE, THE ECONOMICS OF MEDICAL CARE 49-66 (1978).

13. For a discussion of the traditional grounds for governmental intervention in the market, see J. HIRSHLEIFER, *supra* note 11, at 520-51.

necessary.[14] Perhaps the most important example of this need in the health sector is the situation in which, because of publicly funded or tax-subsidized third-party payment plans, the individual participants in an exchange transaction do not bear the full costs associated with the delivery of a service. In such circumstances, some form of governmental intervention may be needed to improve or restore market conditions or incentives. Procompetitive regulations may be required to make the market system function smoothly.

Market proponents, therefore, do not adhere rigidly to the view that no form of governmental regulation is appropriate.[15] They recognize that "the realistic range of public choice does not reflect a total commitment to one or another system of social ordering."[16] They insist, however, that "[i]n a nation whose institutions have relied on market mechanisms for making basic economic choices, governmental imposition of planning [or other forms of coercive regulation] bears a burden of persuasion."[17]

One clear justification for some form of governmental involvement in the medical care marketplace stems from widespread acceptance of the view that society has some responsibility to help needy and worthy persons secure access to medical services. This concern about equity in the distribution of medical care explains the traditional emphasis on access issues in health care discussions. The market system assumes a given distribution of income and does not purport to deal with matters of equity in that distribution. Much governmental intervention in the health sector is designed to further social or redistributive objectives. These goals cannot realistically be fully achieved by the market system on its own because the market system does not focus intrinsically on the redress of perceived distributive inequities.

C. Recognizing the Need to Make Quality/Cost Tradeoffs

Advocates of pro-market policies (marketeers)[18] assume that health is an instrumental, not an ultimate, value.[19] If health is defined

14. For example, improving the flow of information through truth-in-labeling legislation may improve the functioning of the market by promoting informed decisionmaking. Antitrust enforcement is another example of a market-perfecting form of government intervention designed to prohibit anticompetitive conduct and permit the market to function properly. Professors Alain Enthoven and Clark Havighurst, among others, have discussed the potential compatibility of certain forms of regulation and competition. *See* A. ENTHOVEN, *supra* note 7, at 78-82; Havighurst, *Health Planning for Regulation: Implementing the 1979 Regulations*, 44 LAW & CONTEMP. PROBS., Winter 1981, at 33; Havighurst, *Prospects for Competition Under Health Planning-cum-Regulation*, in NATIONAL HEALTH INSURANCE: WHAT NOW, WHAT LATER, WHAT NEVER? 329 (M. Pauly ed. 1980). *See also* Blumstein & Sloan, *Redefining Government's Role in Health Care: Is A Dose of Competition What the Doctor Should Order?*, 34 VAND. L. REV. 849, 854-55, 888-94 (1981) [hereinafter cited as *Redefining Government's Role*].

15. *See generally* Bovbjerg, *Competition Versus Regulation in Medical Care: An Overdrawn Dichotomy*, 34 VAND. L. REV. 965 (1981).

16. *Redefining Government's Role*, *supra* note 14, at 854.

17. Blumstein & Sloan, *Health Planning and Regulation Through Certificate of Need: An Overview*, 1978 UTAH L. REV. 3, 3 [hereinafter cited as *Health Planning and Regulation*].

18. As is indicated in previous writings, the author is a proponent of market-perfecting policies in the health sector. *See, e.g., Redefining Government's Role*, *supra* note 14. The generic term is used to indicate that the views of pro-market advocates are not exclusively those of this author, nor are they necessarily monolithic. They reflect a set of contending viewpoints in the health policy dialogue. *Compare* Sigelman, *Palm Reading the Invisible Hand: A Critical Examination of Pro-Competitive Reform Proposals*, 6 J. HEALTH POL., POL'Y, & L. 578 (1982), *with* Blumstein, *On*

so broadly as to contemplate the positive notion of complete physical, social, and mental well-being, and not merely the absence of disease, it is hard to distinguish the idea of health from that of the general welfare. This, in turn, suggests that health, so broadly defined, is an end in and of itself—"if you have your health, you have everything." This "Geritol" viewpoint makes thinking in terms of economic tradeoffs extremely difficult; if health is an ultimate value, tradeoffs make no sense.

The breadth of the definition of health is not an idle quibble. If we accept the common assumption that people are not responsible for their poor health and the concomitant assumption that society is more munificent in providing financial assistance for the unhealthy than for the poor, who are perceived as less worthy, then the critical political importance of the definition of health becomes apparent. Income redistribution is considerably facilitated if one portrays the beneficiaries of public support as unhealthy, and therefore not culpable, rather than merely as poor, and therefore potentially blameworthy. The rhetoric used in public debate in many areas of public policy reflects the political reality of this subtle distinction. For example, during the era of the Great Society those seeking additional governmental assistance for urban development continually referred to the problem of the nation's "sick" cities, which were in need of help. This physiological metaphor is often used as a rhetorical device in health policy debates. The scope of the definition of health is hardly a value-neutral matter. The implicit assumption seems to be that, with a broader concept of health, redistributive efforts have a better chance of political success.

For market proponents, a narrower definition of health as an instrumental value is very important because it opens up the possibility of thinking seriously about tradeoffs between medical expenditures and spending on other items of consumption. A major attraction of the market system is that private, decentralized decisionmakers are left free to choose among available alternatives. Market proponents assume that "given the opportunity to choose between alternatives and the ability to benefit tangibly from forgoing the consumption of medical services, many consumers (or, when appropriate, physicians on behalf of their patients) will decide to forgo some medical services that would be marginally beneficial because other, nonmedical priorities will take precedence."[20]

The ability of private decisionmakers to benefit from choosing not to consume an item is critical to the effective functioning of this system of private choice.[21] Given a certain budget, decisionmakers in the private market face "either/or" rather than "yes/no" choices—selecting

the Other Hand: Some Comments on Sigelman's Palmistry, 6 J. HEALTH POL., POL'Y, & L. 621 (1982). *See generally Competition and Regulation in Health Care Markets*, 59 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 106 (1981). For a discussion of different perspectives on health policy issues, see R. ALFORD, HEALTH CARE POLITICS 1-21 (1975). *See also* A. DONABEDIAN, ASPECTS OF MEDICAL CARE ADMINISTRATION 3-27 (1973).

19. *See* Zubkoff & Blumstein, *supra* note 10, at 73-74.

20. *Redefining Government's Role*, *supra* note 14, at 858.

21. *See* A. ENTHOVEN, *supra* note 7, at 16; H. FRECH & P. GINSBURG, PUBLIC INSURANCE IN PRIVATE MEDICAL MARKETS 21-22 (1978).

one good or service means foregoing another. However, private decisionmakers are in a good position to make efficient tradeoff decisions based on their own distinctive sets of preferences only when they can "recapture and reallocate (R & R) resources derived from economizing."[22] Without R & R, decisions are seen not as "either/or" but as "yes/no," and a private decisionmaker will be prone to say "yes" to a given consumption opportunity because he will not gain anything by saying "no" to any outlay of funds.

D. The Vision of Equity in a Market System

The instrumental nature of health and the need to allow for R & R by decisionmakers are essential points for market proponents. In addition, they assume that different people, even if similarly situated, will have different consumer preferences.[23] This assumption, coupled with a system of freedom of choice, suggests another important, albeit controversial, conclusion—that total equality in the actual consumption of medical services is not necessary, practical, or even desirable. Given the liberty to choose and different preferences among different consumers, varying levels of consumption are virtually inevitable and not inherently undesirable. Because this view explicitly contemplates, and even advocates, a dual or multitier system of medical care,[24] it is probably most troublesome to traditional access egalitarians, whose ideal is universal access, regardless of ability to pay, to "mainstream" health services.[25]

In contrast to the position espoused by access egalitarians, market proponents recognize that society's commitment to providing access to medical services extends only to those services that are somehow different from other consumer items.[26] Beyond some level of care, variously defined in such terms as "minimal," "basic," "acceptable," or "adequate," the case for further governmental equalization becomes far less compelling.[27] "The competitive approach . . . does not directly address the question of equity of access, except that it rejects the ideological stances that (1) health is an ultimate good and therefore not subject

---

22. *Redefining Government's Role*, *supra* note 14, at 858.

23. *See* E. Koos, The Health of Regionville 30-39 (1954).

24. *See* Blumstein & Zubkoff, *Perspectives on Government Policy in the Health Sector*, 51 Milbank Memorial Fund Q.: Health & Soc'y 395, 407-12 (1973) [hereinafter cited as *Perspectives on Government Policy*].

25. *See* Rosenblatt, *Health Care, Markets, and Democratic Values*, 34 Vand. L. Rev. 1067 (1981). Professor Rosenblatt notes that

> [d]uring the 20th century, most health care reformers apparently have agreed upon the general ideal of universal access—regardless of ability to pay—to a broadly adequate package of socially financed benefits delivered under a single standard that is termed 'mainstream' quality of care. . . . The market approach not only accepts these constraints; it affirmatively *encourages* them, claiming that they are consistent with a redefined ideal. The market advocates, therefore, do not view the ethical ideal of health care delivery as universal access to *broadly* adequate, mainstream benefits, but rather as universal access to *minimally* adequate necessary benefits, coupled with *differential* access to most health care based upon ability to pay.

Id. at 1112-13.

26. *See Perspectives on Government Policy*, *supra* note 24, at 411. *See also* Fried, *Equality and Rights in Medical Care*, Hastings Center Rep., Feb. 1976, at 29.

27. *See* Blumstein & Zubkoff, *Public Choice in Health: Problems, Politics and Perspectives on Formulating National Health Policy*, 4 J. Health Pol., Pol'y, & L. 382, 383-87, 403-05 (1979) [hereinafter cited as *Public Choice in Health*].

to tradeoffs and (2) equal utilization of medical services is an absolute equitable imperative, unrelated to private preferences and different private resource constraints."[28]

E. Institutional Biases Toward the Excessive Allocation of Resources to Medical Services

Although market proponents acknowledge the desirability of providing access to medical care to defined groups of needy persons, they also recognize the institutional risks posed by increasing governmental involvement in health matters. Distributive injustice is not a problem unique to or caused by the health care system; this injustice stems not from "an inability to perform the medical delivery function well but [from] an inability to transfer income in the process."[29] Because distributive concerns are not limited to the health field and do not derive from unique characteristics of the health sector, market proponents who are sensitive to the institutional impact of redistributive programs on the functioning of the marketplace would seek to minimize the distorting effect of governmental redistributive efforts on private health care decisionmaking.

1. Health As Symbolism.—Explicit identification of government with the provision of lifesaving medical services raises important symbolic issues. It is very difficult for a democratic government to assume responsibility for directly harming or refusing to help someone whom it is perceived as being obligated to assist. A government must decide whether to provide every medical service, even those of small benefit to its beneficiaries, or to limit its outlays in some way. If it decides to restrict these outlays, the problem of establishing acceptable allocative criteria arises; rationing in the generic sense becomes a relevant consideration.

These allocative decisions are constrained by political symbolism. The societal notion that life is priceless may be comforting, but upon close analysis it turns out to be a myth. Individuals and governments often act in ways that belie the "life is priceless" dictum. People ignore speed limits, refuse to wear seat belts, eat unhealthy foods, drink diet soda, smoke, climb mountains, sky dive, and engage in numerous other hazardous activities. Life and health are risked for a variety of reasons, ranging from economic reward to leisure time excitement. Yet, when government, as the representative of society, focuses on the life or health of a single person under its wing, it feels obliged to act in accordance with the myth that life is priceless.[30] "Society's humanitarian self-image may be at stake, and we may be willing to expend considera-

28. *Redefining Government's Role*, *supra* note 14, at 867.

29. *Public Choice in Health*, *supra* note 27, at 383.

30. *See* Zeckhauser, *Coverage for Catastrophic Illness*, 21 PUB. POL'Y 149 (1973); Zeckhauser, *Procedures for Valuing Lives*, 23 PUB. POL'Y 419 (1975). *See also* Fried, *The Value of Life*, 82 HARV. L. REV. 1415 (1969); Fuller, *The Case of the Speluncean Explorers*, 62 HARV. L. REV. 616, 623 (1949); Schelling, *The Life You Save May Be Your Own*, in PROBLEMS IN PUBLIC EXPENDITURE ANALYSIS 127 (S. Chase ed. 1968).

ble sums of money, ostensibly to save a life, but also to preserve a valuable myth."[31]

A consideration of the institutional arrangements that govern allocative decisions is important, because the process of decisionmaking can have a significant impact on the outcome of aggregate resource allocation decisions. A substantial risk of symbolic blackmail exists;[32] "there will be pressure to keep up appearances by allocating more and more resources which might not otherwise be so spent, since there seems to be an independent value assigned to maintenance of the humanitarian image itself."[33]

2. Identifiable and Statistical Lives.—The existence of the "lifesaving imperative"[34] is most acute "whenever government is obviously placed in a position to undertake a specific lifesaving effort."[35] Where government's role is more attenuated, "the destructive symbolic effect of its identification with specific human suffering" is easier to avoid.[36] Thus, it is useful to draw a distinction between government's role in dealing with *identifiable* lives and *statistical* lives.[37]

The lifesaving imperative is at its strongest when governmental decisionmaking overtly and explicitly affects an identifiable individual for whom the government is perceived to be responsible. In such circumstances, it is very difficult to restrict governmental munificence. A poignant illustration of this phenomenon is the dilemma of capital punishment. Regardless of one's views on the capital punishment issue, reaching an ultimate resolution is not an easy matter. This stems mainly from the unavoidably direct role government plays in putting a person to death. Even in the case of a convicted killer unquestionably guilty of a heinous crime, we are very nervous about our government's making an affirmative decision to take a human life. "The more visible and more directly accountable the government becomes" in allocation questions involving identifiable individuals, "the more nearly analogous to capital punishment decisions these issues become."[38] The effect is a subtle form of symbolic blackmail that induces society "to spend more resources in this area than it might otherwise choose."[39]

This analysis suggests that government should consciously avoid becoming directly entangled with micro allocation decisions that affect identifiable lives. Macro allocation decisions, which deal primarily

31. Blumstein, *Constitutional Perspectives on Governmental Decisions Affecting Human Life and Health*, 40 LAW & CONTEMP. PROBS., Autumn 1976, at 231, 233 [hereinafter cited as *Constitutional Perspectives*].

32. *See id.* at 303.

33. *Public Choice in Health*, *supra* note 27, at 389.

34. *See* Havighurst, Blumstein & Bovbjerg, *Strategies in Underwriting the Costs of Catastrophic Disease*, 40 LAW & CONTEMP. PROBS., Autumn 1976, at 122, 140-45.

35. *Id.* at 140.

36. *Id.* at 143.

37. *See generally* Havighurst & Blumstein, *Coping with Quality/Cost Trade-Offs in Medical Care: The Role of PSROs*, 70 NW. U.L. REV. 6, 21-25 (1975).

38. *Constitutional Perspectives*, *supra* note 31, at 252.

39. *Id.* Another example of this phenomenon is the federal financing of kidney dialysis for victims of end-stage renal disease. For a discussion of the enactment of the federal legislation, suggesting that the "myth" explanation tends to be exaggerated, see Rettig, *The Policy Debate on Patient Care Financing for Victims of End-Stage Renal Disease*, 40 LAW & CONTEMP. PROBS., Autumn 1976, at 196.

with statistical lives, might be able to defuse some of the symbolism associated with governmental involvement in the medical sector. Direct public attempts to contain health care costs, however, risk governmental identification with and responsibility for "every delayable death and preventable hurt."[40] Decentralized choices by nongovernmental decisionmakers such as households, labor unions, insurance companies, hospital committees, and physicians have "greater potential for precluding symbolic concerns from becoming inextricably involved in policy formulation and will likely point more attention to necessary economic tradeoffs."[41]

3. The Impact of Third-Party Payments.—In the health sector, collective decisionmaking has become a fact of life. Private third-party health insurance has grown dramatically since World War II,[42] promoted at first by hospitals seeking to assure payment for inpatient treatment and induced by favorable income tax treatment.[43] Corporate contributions for health insurance premiums are deductible as ordinary business expenses, but they are not imputed as taxable income to employees. This provision encourages employees and their union representatives to seek broader and deeper health insurance coverage, since premiums are paid in untaxed dollars.[44] The tax law also encourages unions to carry out a bargaining strategy that conforms to their traditional ideological commitment to improved health benefits. Extensive medical coverage is a highly visible piece of a fringe benefit package, which can easily be described and sold to a union's membership.

Since the advent of public third-party financing programs under Medicare and Medicaid, governmental financial involvement in the medical sector has been direct and substantial.[45] In general, treatments that a physician believes to be medically desirable or necessary are reimbursable.[46] For both patients and their physicians, the pressures for "spare-no-expense" medicine are apparent. Worried patients urge their doctors to provide the best care available. A well-meaning physician can accommodate this understandable instinct quite painlessly because insurance reduces a patient's out-of-pocket expenses to only a fraction of the total costs generated by an episode of illness. The incentive to conserve on the part of patient or provider is minimal.[47]

---

40. Havighurst, Blumstein & Bovbjerg, *supra* note 34, at 157-158.

41. *Public Choice in Health*, *supra* note 27, at 389-90.

42. For example, in 1940 direct consumer payments accounted for 81.3% of total personal health care expenditures, whereas they accounted for only 32.4% in 1980. Similarly, private insurance accounted for only 9.1% of total personal health care expenditures in 1950, whereas it accounted for 26.6% of total personal health care expenditures in 1980. *See* Gibson & Waldo, *supra* note 6, at 38.

43. *See generally* S. Law, Blue Cross: What Went Wrong? (2d ed. 1976).

44. *See* A. Enthoven, *supra* note 7, at 19-21. *See also* Feldstein, *The Medical Economy*, Sci. Am., Sept. 1973, at 151; Havighurst & Blumstein, *supra* note 37, at 13 n.27.

45. During the decade of the 1970s, federal expenditures under the Medicare program more than doubled over each five-year span. In 1970 total federal Medicare expenditures were $7.5 billion. That figure rose to $16.3 billion in 1975 and $36.7 billion in 1980. Similarly, total expenditures for Medicaid rose from $5.5 billion in 1970 to $14.2 billion in 1975 and $26.6 billion in 1980. *See* Gibson & Waldo, *supra* note 6, at 48-52.

46. *See* A. Enthoven, *supra* note 7, at 21-23.

47. *See* Havighurst & Blumstein, *supra* note 37, at 13-15, 27-28, 31-33; *Redefining Government's Role*, *supra* note 14, at 855-59.

The divergence between the costs facing a patient during a visit to a medical care provider and the actual costs incurred during that visit creates a powerful structural bias for the allocation of more resources to medical services than might be the case if patients experienced a closer relationship between benefits accrued and costs incurred.[48] This incentive toward oversupply of medical services has added impetus to the search for mechanisms to control health care costs, which are escalating beyond projections and intruding upon other needs for scarce public and private funds.

III. Constitutional Constraints on Rationing Medical Care

The interest in containing costs in collectively financed programs raises the specter of some form of rationing. The existence of substantial public financing programs through Medicare and Medicaid and of vast private health insurance funds suggests that unconstrained decentralized decisionmaking may not result in an optimal level of health care spending. Many would now question, however, whether regulatory intervention, especially at the micro level, is the best method of achieving optimal resource allocation decisions in medical care. Market-oriented incentive approaches are now receiving the attention they deserve. Nevertheless, these market alternatives have their critics,[49] and governmental intervention in the health policy arena is certain to continue in some form. It is therefore appropriate to consider how legal and constitutional principles constrain both government and private decisionmaking associated with the rationing of medical care.

The distinctions between public and private decisionmaking on the one hand and macro and micro decisionmaking on the other hand are important factors in considering constitutional constraints on governmental rationing decisions. Government can ration the use of public funds allocated to medical care,[50] and it can constrict the use of private funds for the purchase of medical services, equipment, and other medical supplies.[51] Government can also circumscribe the autonomy of medical practitioners in their professional activities.[52] Moreover, rationing can occur at both a macro and a micro level. A macro decision could involve a legislative or administrative judgment affect-

48. *See* Newhouse, Manning, Morris, Orr, Duan, Keeler, Leibowitz, Marquis, Marquis, Phelps & Brook, *Some Interim Results From a Controlled Trial at Cost Sharing in Health Insurance*, 305 NEW ENG. J. MED. 1501 (1981). *See also* H. FRECH & P. GINSBURG, *supra* note 21, at 17 (discussing the concept of "moral hazard").

49. *E.g.*, Rosenblatt, *supra* note 25; Sigelman, *supra* note 18; Young & Saltzman, *Can Competition Really Lower Health Costs*, Wall. St. J., Dec. 22, 1981, at 14, col. 3. *See also* Marmor, Boyer & Greenberg, *Medical Care and Procompetitive Reform*, 34 VAND. L. REV. 1003 (1981).

50. The government can accomplish this in a number of ways. It can reduce the coverage of a program, cap the total expenditures for a program, establish more stringent criteria for eligibility for a program, or eliminate certain types of programs entirely.

51. The most significant example of this role for government is the imposition of certificate-of-need requirements for private investment in facilities and equipment. *See* Havighurst, *Regulation of Health Facilities and Services by "Certificate of Need,"* 59 VA. L. REV. 1143 (1973); text accompanying notes 219-27 *infra*. *See also Health Planning and Regulation*, *supra* note 17, at 23-32; Salkever & Bice, *The Impact of Certificate-of-Need Controls on Hospital Investment*, 54 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 185 (1976). *See generally* D. SALKEVER & T. BICE, HOSPITAL CERTIFICATE-OF-NEED CONTROLS: IMPACT ON INVESTMENT, COSTS, AND USE (1979); F. SLOAN & B. STEINWALD, INSURANCE, REGULATION, AND HOSPITAL COSTS (1980).

52. *See, e.g.*, Whalen v. Roe, 429 U.S. 589 (1977) (statute requiring recordation of prescription holders' identities upheld).

ing coverage or eligibility under a governmental program,[53] or a governmental choice to sponsor research on cancer but not Tay-Sachs disease. An example of a micro decision would be the determination of whether a person broadly eligible for benefits under a public program qualifies for treatment in a given set of circumstances.[54] This macro/micro distinction is particularly important in the procedural due process area.

When government rations public funds, disappointed claimants can allege that the government's action is impermissibly discriminatory and denies them equal protection of the laws.[55] In the alternative, they can claim that government is breaching some affirmative duty toward them in such a way as to deny them life, liberty, or property under the due process provisions of the fifth or fourteenth amendment. When government restricts a private person's ability to spend funds on medical services, the likely claim would again be that government is overreaching in an area protected as life, liberty, or property without due process of law. In the discussion that follows, these potential complaints will be considered.

A. Rationing Public Funds: Procedural Due Process

The due process provisions of the fifth and fourteenth amendments have both procedural and substantive components. This subsection will focus on the procedural aspect of due process, and the next subsection will consider the substantive component. In either situation, a threshold issue is the existence of state action.

1. State Action.—In order for the provisions of the fifth and fourteenth amendments to be applicable at all, the decisions at issue must either be made by governmental bodies or by authorities whose actions can be fairly attributable to the government.[56] Purely private rationing decisions made by physicians in their everyday practice of medicine are not likely to be attributed to the state. Similarly, a rationing decision by a private health maintenance organization (HMO) would probably be deemed outside the ambit of constitutional protection. The marketplace, informal peer review and referral networks, formal professional disciplinary bodies, and the law of malpractice must be relied on to constrain such private allocative decisions.

Government is involved in the medical delivery system in a number of ways, however, and in such circumstances the constraints of

53. An example of a coverage decision is the determination in a number of states to reduce the number of hospital days covered under Medicaid programs. These cutbacks have been part of a general cost-containment strategy necessitated by the reduction of public funds available for government health programs. *See* Jennings v. Alexander, 518 F. Supp. 877 (M.D. Tenn. 1981) (state's proposed reduction of Medicaid coverage does not violate federal law).

54. Preterm, Inc. v. Dukakis, 591 F.2d 121, 125 (1st Cir. 1979).

55. Although the equal protection clause does not directly apply to federal action, the Supreme Court has construed the due process language of the fifth amendment as imposing a duty on the federal government to guarantee equal protection of the laws, just as the fourteenth amendment establishes such a constraint on the actions of state and local governments. Bolling v. Sharpe, 347 U.S. 497 (1954).

56. *See* United States v. Classic, 313 U.S. 299 (1941) (election officials who altered ballots acted under color of state law).

the Constitution apply. For example, the government finances medical care directly through Medicare and Medicaid. These programs involve both the establishment and the application of eligibility criteria for medical care. In some cases governmental entities, such as state and local hospitals, are involved in the direct provision of care to defined populations. More indirectly, government provides funds for use by private groups that then supply medical services to specific client groups. Examples of such funding include federal support for hospital construction through the Hill-Burton Program and federal grants for a variety of agencies that furnish services to specific groups of beneficiaries.

When the government directly allocates public funds or rations medical care, the state action threshold is easily crossed. Government officials acting in their official capacities are likewise typically viewed as acting under the color of state law when establishing criteria of program eligibility or making determinations concerning individual qualifications in specific cases.[57] A more difficult question arises at the lower stages of decisionmaking when administrative details for micro allocation are left to private persons or institutions.

In light of the discussion in the previous section, the government rationally might seek to distance itself from close public identification with the nitty-gritty problems of micro allocation. Heeding the admonition that direct governmental involvement is an open invitation to symbolic blackmail, the government could choose to allocate sums of money in the macro sense, with only general guidelines. By leaving the details of micro allocation to private actors, it could skirt the ticklish matter of centrally dictating allocation criteria. The question then becomes whether these private persons or institutions are operating under the color of state law, which would subject their behavior to constitutional scrutiny.[58]

The Supreme Court has sought for years to draw the line between governmental and private responsibility. Much of the state action doctrine has developed from the civil rights litigation of the past three decades.[59] Before the enactment of the major civil rights legislation of the 1960s, aggrieved blacks could not sustain claims of discrimination unless they somehow demonstrated governmental responsibility for the allegedly discriminatory acts. Constitutional guarantees were inapplicable in the absence of state action, and no specific civil rights legislation yet existed to eliminate private forms of discrimination in public accommodations, employment, and housing. In light of that reality and in conformity with the Warren Court's philosophy that government should bear the responsibility for a greater part of the nation's

57. The federal civil rights statute, 42 U.S.C. § 1983 (1976), provides relief against those who act under color of state law to deprive a person of constitutional rights. In determining whether state action exists under this provision, some courts have held that action by the federal government is irrelevant, since § 1983 applies only to state officials. *See* Blackburn v. Fisk Univ., 443 F.2d 121, 123 (6th Cir. 1971).

58. *Constitutional Perspectives*, *supra* note 31, at 240.

59. *See, e.g.,* Peterson v. Greenville, 373 U.S. 244 (1963); Lombard v. Louisiana, 373 U.S. 267 (1963); Burton v. Wilmington Parking Auth., 365 U.S. 715 (1961).

social problems, the Supreme Court endorsed an expansive state action principle.[60]

More recently, however, the Supreme Court has backed away from this far-reaching approach. The Court has become increasingly willing to define behavior as private activity and, therefore, not subject to constitutional limitations.[61] This judicial trend indicates that government now may well be able to implement successfully the strategy of keeping distance between itself and the ultimate allocative decisions in medical care, which might be perceived as involving life or death. The Warren Court might have thwarted this technique by finding a perceived affirmative duty to perform a "governmental function,"[62] but more recent cases suggest the present Court's disinclination to accept, much less extend, such positivistic notions.[63]

Traditionally, state action analysis has had two dimensions. On the one hand, the nature and extent of governmental participation in and support of a private party's conduct has justified a finding of state action.[64] The other strand of the doctrine held private action that performed a governmental or public function attributable to the state.[65] In the last ten years the Court has narrowed the state participation theory[66] and virtually eviscerated the public function theory.[67]

A good illustration of the type of case that can raise very difficult state action issues involves the federal government's Hill-Burton health facilities construction program.[68] The Hill-Burton program granted federal funds to private hospitals for new construction. The recipients of these funds were obligated to provide a reasonable amount of medical care to persons unable to pay.[69] When indigents and individual employees began to challenge various activities of the private health facilities on constitutional grounds,[70] each challenge faced the threshold question of whether the receipt of federal construction funds by the hospital transformed its conduct into state action subject to the limitations of the Constitution.

In the context of employment litigation (where a doctor seeks staff privileges, for example), many federal courts have held that the receipt of Hill-Burton funds is by itself insufficient to transmute all hospital

60. *Compare* Polk County v. Dodson, 102 S. Ct. 445 (1981), *with* Shelley v. Kraemer, 334 U.S. 1 (1948).

61. *See generally* Pasadena City Bd. of Educ. v. Spangler, 427 U.S. 424 (1976); Jackson v. Metropolitan Edison Co., 419 U.S. 345 (1974); Moose Lodge No. 107 v. Irvis, 407 U.S. 163 (1972); Evans v. Abney, 396 U.S. 435 (1970).

62. *See generally* Evans v. Newton, 382 U.S. 296 (1966).

63. *See, e.g.,* Flagg Bros. v. Brooks, 436 U.S. 149 (1978); Hudgens v. NLRB, 424 U.S. 507 (1976).

64. *See* Burton v. Wilmington Parking Auth., 365 U.S. 715 (1961).

65. *See* Marsh v. Alabama, 326 U.S. 501 (1946).

66. *See* Jackson v. Metropolitan Edison Co., 419 U.S. 345 (1974); Moose Lodge No. 107 v. Irvis, 407 U.S. 163 (1972).

67. *See* Flagg Bros. v. Brooks, 436 U.S. 149 (1978).

68. For a description and analysis of the Hill-Burton program, see J. LAVE & L. LAVE, THE HOSPITAL CONSTRUCTION ACT: AN EVALUATION OF THE HILL-BURTON PROGRAM, 1948-1973 (1974); Note, *The Hill-Burton Act, 1946-1980: Asynchrony in the Delivery of Health Care to the Poor*, 39 MD. L. REV. 316 (1979).

69. *See* Note, *Due Process for Hill-Burton Assisted Facilities*, 32 VAND. L. REV. 1469 (1979).

70. *E.g.,* Newsom v. Vanderbilt Univ., 453 F. Supp. 401 (M.D. Tenn. 1978), *aff'd in part, modified in part, and rev'd in part*, 653 F.2d 1100 (6th Cir. 1981).

conduct into state action.[71] The courts have required both direct governmental participation in the challenged conduct and a relationship between the federal funding and the challenged conduct.[72]

The attribution of state action to Hill-Burton-funded hospitals is a much closer question when indigent plaintiffs challenge a hospital's fulfillment of its uncompensated care obligation. In its recent decision in *Newsom v. Vanderbilt University*,[73] the Sixth Circuit had little difficulty disposing of the plaintiffs' "public function" theory. Citing recent definitive Supreme Court precedent on the public function question, the court concluded that the doctrine "only applies to those functions which were *exclusively* reserved to the states traditionally."[74] Since states have not traditionally reserved to themselves the function of providing hospital services, the court found the public function theory inapplicable.

The more troublesome state action question in *Newsom* was whether federal funding of the hospital's construction and the hospital's assurance that it would provide a reasonable volume of care to those unable to pay meant that its administration and implementation of the Hill-Burton program were attributable to the state. The court did not reach this issue, but its opinion recognized the difficulty caused by federal and state regulations that left a great deal of discretion to the hospital. Plaintiffs' advocates naturally saw this as a shirking of public responsibility, a view accepted by the trial judge in *Newsom*. The trial court held that the hospital's method of allocating its Hill-Burton uncompensated care "should be devised" under governmental direction.[75] The facts showed almost no governmental involvement in the establishment of rationing criteria, and the hospital apparently did not set up identifiable standards. Indeed, the failure of government to act in the situation was precisely the cause of the plaintiffs' grievance: "Neither the State nor federal [government] has issued any regulation or made any other effort to require . . . any . . . Hill-Burton subsidized health care facility in Tennessee to allocate the required volume of free or below-cost services according to fair and orderly procedures . . . ."[76] Government inaction rather than action apparently posed the problem from the plaintiffs' point of view.

The nature of this complaint suggests that the representatives of the indigent plaintiffs in *Newsom* were arguing for an affirmative obli-

71. Aasum v. Good Samaritan Hosp., 542 F.2d 792 (9th Cir. 1976); Briscoe v. Bock, 540 F.2d 392 (8th Cir. 1976); Jackson v. Norton-Children's Hosps., Inc., 487 F.2d 502 (6th Cir. 1973), *cert. denied*, 416 U.S. 1000 (1974); Ward v. St. Anthony Hosp., 476 F.2d 671 (10th Cir. 1973); Sament v. Hahnemann Medical College & Hosp., 413 F. Supp. 434 (E.D. Pa. 1976); Barrett v. United Hosp., 376 F. Supp. 791 (S.D.N.Y. 1974). The sole exception is the Fourth Circuit, which has found the receipt of Hill-Burton funds to constitute state action. Simkins v. Moses H. Cone Memorial Hosp., 323 F.2d 959 (4th Cir. 1963), *cert. denied*, 376 U.S. 938 (1964). For a general discussion, see Cronin, *Private Hospitals That Receive Public Funds Under the Hill-Burton Program: The State Action Implications*, 12 NEW ENG. L. REV. 525 (1977).

72. *See generally* McCoy, *Current State Action Theories, The* Jackson *Nexus Requirement, and Employee Discharges by Semi-Public and State-Aided Institutions*, 31 VAND. L. REV. 785 (1978).

73. 653 F.2d 1100 (6th Cir. 1981).

74. *Id.* at 1114 (emphasis in original) (citing Flagg Bros. v. Brooks, 436 U.S. 149 (1978)).

75. Newsom v. Vanderbilt Univ., 453 F. Supp. 401, 420 (M.D. Tenn. 1978), *aff'd in part, modified in part, and rev'd in part*, 653 F.2d 1100 (6th Cir. 1981).

76. *See* Joint Appendix at 27.

gation on the part of government to involve itself in specific allocative decisions whenever it funds a program at a more general level. Government had not explicitly, or even implicitly, approved the private hospital's action. That the state had not ordered the hospital to allocate its Hill-Burton care in any particular manner was particularly significant in light of the Supreme Court's explanation that a "State is responsible for the . . . act of a private party when the State, by its law, has compelled the act."[77] In *Newsom* there was no imprimatur of governmental action and no formal acquiescence. In short, "the crux of [the *Newsom*] complaint [was] not that the state *ha[d]* acted, but that it ha[d] *refused* to act."[78]

If the discussion in the first part of this Article is taken seriously, government should seek to avoid the specter of "openly renouncing the widely held belief in the sanctity of human life, acknowledging for all to see that society is unwilling to expend the funds necessary to obviate the scarcity"[79] in resources that was an inherent feature of the Hill-Burton program's basic design.[80] The courts should, as seems likely in light of recent Supreme Court trends[81] and the discussion in *Newsom*, reject the affirmative responsibility notion. If they do not, the strategy of distancing government from certain basic allocative decisions may run into constitutional headwinds. The courts have been and should remain very cautious in developing positivistic notions of an affirmative duty to provide medical care. Prudence is particularly warranted when a finding of state action would impose procedural requirements that could amount to "legalistic indecency."[82]

While avoidance of constitutional scrutiny by skirting the state action requirement does have its attraction, it also creates some risks. Since the existence of state action is a prerequisite for any constitutional scrutiny under either due process or equal protection, a finding of no state action insulates conduct from judicial review under constitutional standards, irrespective of the substantive character of the claim. Some substantive outcomes are so unacceptable that their social costs would exceed the benefits of decentralization and governmental distancing. For instance, although decentralized decisionmaking necessarily contemplates some inconsistency and perceived unfairness in individual cases, it also threatens to permit patterns of racial discrimination to remain constitutionally unreviewable.

Fortunately, there are two responses to these concerns. First, the monolithic character of the state action doctrine seems to be gradually eroding. To some extent, the courts now recognize varying degrees of governmental responsibility in different contexts. Thus, some circuit courts of appeal have in practice, if not in theory, applied different

77. Adickes v. Kress & Co., 398 U.S. 144, 170 (1970).

78. Flagg Bros. v. Brooks, 436 U.S. 149, 166 (1978) (emphasis in original).

79. *Constitutional Perspectives*, *supra* note 31, at 250.

80. The Hill-Burton program required hospitals to provide a reasonable volume of unreimbursed services to indigent patients. It did not create a universal entitlement to free medical care for all indigent patients.

81. *See, e.g.*, Flagg Bros. v. Brooks, 436 U.S. 149 (1978).

82. *See* Havighurst, Blumstein & Bovbjerg, *supra* note 34, at 157.

standards of state action where racial discrimination is involved.[83] Second, the extensive civil rights legislation of the 1960s in such areas as public accommodations, housing, employment, and use of federal funds makes most private discrimination in the delivery of medical services illegal.[84] Title VI of the 1964 Civil Rights Act expressly bars discrimination in the use of federal funds, and the legislative standard is at least as stringent as the constitutional.[85] Title VI thus provides a legislative remedy against the most troublesome potential abuses. At the same time, it allows government to distance itself from sensitive rationing choices that might best be left to nongovernmental decisionmakers. The civil rights law adequately checks potential discriminatory use of federal funds.

Other than constitutional and statutory constraints, a hospital faces both public pressure not to make decisions arbitrarily and legal attack for tortious conduct should it breach a duty to a patient. A hospital staff's awareness of its potential civil responsibility diminishes the likelihood of unacceptably egregious conduct.[86] Thus, other social and legal controls largely obviate the need for the application of constitutional safeguards in the micro allocation process. Unless government affirmatively chooses to become a more active participant in specific allocative decisions, the courts can comfortably apply the state action doctrine in accordance with the recent trend of restraint. The lack of a constitutional remedy for putative procedural due process violations does not appear to pose a serious threat to basic societal interests, and the ability of government to avoid the direct confrontation of values inherent in the process of developing explicit allocative criteria should allow nongovernmental institutions to manage the rationing problem with a minimum of demoralization.

2. Deprivation of Constitutionally Protected Interests.—Once a court finds conduct attributable to the government, it triggers the safeguards of the fifth and fourteenth amendments.[87] The procedural protections of due process apply only when deprivation of a life, liberty, or property interest within the meaning of the fifth and fourteenth amendments has occurred. At the threshold, "to determine whether due process requirements apply in the first place, we must look not to the 'weight' but to the nature of the interest at stake."[88] Interests protected by the due process clause are those to which a person has a legitimate claim of legal entitlement, not a mere unilateral expectation.[89] In the seminal decision *Board of Regents v. Roth*,[90] and in subsequent decisions, the Supreme Court has made it clear that protected property in-

83. *See generally* Black, *The Supreme Court, 1966 Term—Foreword: "State Action," Equal Protection, and California's Proposition 14*, 81 HARV. L. REV. 69 (1967); *Constitutional Perspectives*, *supra* note 31, at 240 n.44.

84. *See* Havighurst, Blumstein & Bovbjerg, *supra* note 34, at 164-65 & n.186.

85. *See* Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978) (plurality opinion).

86. *See, e.g.*, Darling v. Charleston Community Memorial Hosp., 33 Ill. 2d 326, 211 N.E.2d 253 (1965), *cert. denied*, 383 U.S. 946 (1966).

87. *See generally* Van Alstyne, *Cracks in "The New Property": Adjudicative Due Process in the Administrative State*, 62 CORNELL L. REV. 445 (1977).

88. *See* Board of Regents v. Roth, 408 U.S. 564, 570-71 (1972).

89. *See id.* at 577.

90. 408 U.S. 564 (1972).

terests are established by reference to statutes, regulations, administrative practices, contractual arrangements, and other forms of mutual understanding, not by the Constitution itself. In each case an individual must have a legally cognizable and enforceable interest. The protection of procedural due process is designed to safeguard individuals against arbitrary governmental action that "seeks to remove or significantly alter that protected status."[91]

Roth is a good analytical starting point to determine what types of interests qualify for due process protection. In *Roth* an untenured college professor claimed that he was entitled to notice and a hearing to dispute nonrenewal of his contract. The Supreme Court held that the state had discretion to decide whether to retain Roth without providing a hearing or other aspects of due process, because Professor Roth had no legally enforceable claim to reemployment after the expiration of his contract. He was, of course, free to apply for renewal, just as any applicant who is eligible for a job may apply for it. However, the fact that he met the minimum qualifications for the position was not sufficient to give rise to a protected entitlement. It created more of a unilateral hope or expectation, to which due process guarantees are not applicable.

A series of prison cases further illustrates the difference between a unilateral expectation and a legitimate claim of entitlement. In *Meachum v. Fano*,[92] the Court rejected a prisoner's claim of entitlement to a hearing when transferred from one prison to another, concluding that no entitlement to due process protection exists as long as governmental decisionmakers have discretion to make a judgment without the application of definitive, objective standards.[93] The Supreme Court in *Greenholtz v. Inmates of the Nebraska Penal and Correctional Complex*[94] also refused to apply due process protections to parole-release determinations, because no entitlement exists to parole from a valid prison sentence unless a statute explicitly confers such an entitlement.[95]

The Court has drawn a distinction between cases involving parole release and those involving parole-revocation determinations. In *Morrissey v. Brewer*,[96] the Court had held that a parole-revocation determination must meet due process standards. The difference between parole release and parole revocation stems from the "nature of the decision that must be made in each case."[97] Parole-revocation determinations involve a two-step process. First, a factfinder must decide whether a condition of parole has been violated. Second, it must deter-

91. Paul v. Davis, 424 U.S. 693, 711 (1976).

92. 427 U.S. 215 (1976).

93. No state law conferred a right on a prisoner to remain in the prison to which he was assigned. The Court concluded that "[w]hatever expectation the prisoner may have in remaining in a particular prison so long as he behaves himself, it is too ephemeral and insubstantial to trigger procedural due process as long as prison officials have discretion to transfer him for whatever reason or for no reason at all." *Id.* at 229.

94. 442 U.S. 1 (1979).

95. The Court pointed out that "[t]here is no constitutional or inherent right of a convicted person to be conditionally released before the expiration of a valid sentence." *Id.* at 7; *accord,* Connecticut Bd. of Pardons v. Dumschat, 452 U.S. 458 (1981).

96. 408 U.S. 471 (1972).

97. *Greenholtz*, 442 U.S. at 9.

mine whether a parolee should be recommitted, either for the benefit of the parolee himself or for the benefit of society. Whereas the second decision requires the exercise of judgment and discretion, the first determination is "a wholly retrospective factual question."[98] If a parolee can show definitively that a condition of parole was not violated, he will not be subjected to parole revocation; he has a legal right to be at liberty, and the state has no basis for revoking parole. There must be a finding of cause to justify parole revocation.[99]

The Court in *Greenholtz* concluded that the parole-release determination "is more subtle" since it depends on an "amalgam of elements, some of which are factual but many of which are purely subjective," involving the "sensitive task of evaluating the advisability of parole release."[100] Crucial to understanding the concept of entitlement is the *Greenholtz* Court's statement that in a parole-release situation, "there is no set of facts which, if shown, mandate a decision favorable to the individual."[101] The parole board's decision turns on a "discretionary assessment of a multiplicity of imponderables . . . ."[102] In sum, in the parole-release situation the state holds out only the *"possibility"* of parole, which is "no more than a mere hope that the benefit will be obtained,"[103] and thus not an entitlement protected by due process.

Connecticut Board of Pardons v. Dumschat,[104] a due process case involving an application for commutation of a life prison sentence, reaffirmed the message of *Greenholtz*. As a matter of practice, the Connecticut Board of Pardons granted about seventy-five percent of the applications for commutation of life sentences. The plaintiff alleged that this pattern of decisionmaking created an entitlement protected by due process. Noting that a determination to commute a long-term sentence "generally depends not simply on objective factfinding, but also on purely subjective evaluations and predictions of future behavior,"[105] the Court concluded that a commutation decision was similar in character to a parole-release determination. As for the practice of granting commutations to most life inmates, the Court held that "[a] constitutional entitlement cannot 'be created . . . merely because a wholly and *expressly* discretionary state privilege has been granted generously in the past.' No matter how frequently a particular form of clemency has been granted, the statistical probabilities standing alone generate no constitutional protections . . . ."[106]

The cases also suggest that the due process clause protects only the "present enjoyment" of an entitlement. Thus, in *Roth* the Court observed that "[t]he Fourteenth Amendment's procedural protection of

98. Morrissey v. Brewer, 408 U.S. at 479.
99. *Id.* at 479-80.
100. 442 U.S. at 9-10.
101. *Id.* at 10.
102. Kadish, *The Advocate and the Expert—Counsel in the Peno-Correctional Process*, 45 MINN. L. REV. 803 (1961).
103. 442 U.S. at 11.
104. 452 U.S. 458 (1981).
105. *Id.* at 464.
106. *Id.* at 465 (quoting Leis v. Flynt, 439 U.S. 438, 444 n.5 (1979) (emphasis in original)); *accord,* Jago v. Van Curen, 102 S. Ct. 31 (1981) (per curiam).

property is a safeguard of the security of interests that a person *has already* acquired in specific benefits."[107] Similarly, in *Paul v. Davis*,[108] which rejected a claimed deprivation of a person's "liberty" interest in his reputation, the Court noted that state law did not grant "any legal guarantee of *present enjoyment* of reputation"[109] that was impaired by the action of public officials. And in *Greenholtz* the Court observed that "[t]here is a crucial distinction between being deprived of a liberty one has, as in parole, and being denied a conditional liberty that one desires."[110] Quoting approvingly from Judge Henry Friendly, Chief Justice Burger agreed that "there is a human difference between losing what one has and not getting what one wants."[111]

The Court, in fact, has found a due process entitlement to exist only in cases involving revocation of existing, presently enjoyed benefits.[112] For example, in *Goldberg v. Kelly*,[113] a landmark decision that has been described as "launching the procedural due process revolution of the early 1970s,"[114] the Court required a due process hearing before the termination of welfare benefits. Thus, the "present enjoyment" notion seems to limit the concepts of liberty and property, but the scope of the limitation will be unclear until the Court explicitly states whether it is a necessary component for a finding of a protected interest.

To recapitulate, for a protected interest to exist, a set of objective facts must exist which, if shown, automatically mandate a favorable decision for the plaintiff. Also, a beneficiary might have to be currently enjoying an interest before he can claim due process protection for it.

The due process cases suggest that the legislative and executive branches of government are capable of making broad allocative choices distributing or rationing medical resources free from procedural due process scrutiny. Despite the imaginative and valiant efforts of commentators such as Professor Rand Rosenblatt[115] to suggest a structural due process requirement even for such macro allocation decisions, the courts have not accepted that position. Thus, due process is not typically required when governmental institutions act in a legislative or a policymaking capacity.[116] Hearings required in these contexts are usually imposed as a result of a statutory or administrative provision, not as a matter of due process.[117]

---

107. 408 U.S. at 576 (emphasis added).
108. 424 U.S. 693 (1976).
109. *Id.* at 711-12 (emphasis added).
110. 442 U.S. at 9.
111. *Id.* (quoting Friendly, *Some Kind of a Hearing*, 123 U. PA. L. REV. 1267, 1296 (1975)).
112. *See, e.g.*, Goss v. Lopez, 419 U.S. 565 (1975); Bell v. Burson, 402 U.S. 535 (1971).
113. 397 U.S. 254 (1970).
114. G. GUNTHER, CONSTITUTIONAL LAW 647 (10th ed. 1980).
115. *See* Rosenblatt, *Health Care Reform and Administrative Law: A Structural Approach*, 88 YALE L.J. 243 (1978).
116. *See* City of Eastlake v. Forest City Enter., Inc., 426 U.S. 668 (1976); Bi-Metallic Inv. Co. v. State Bd. of Equalization, 239 U.S. 441 (1915).
117. *See* Benton v. Rhodes, 586 F.2d 1 (6th Cir. 1978). *See also* Budnicki v. Beal, 450 F. Supp. 546, 555-56 (E.D. Pa. 1978). An illustration of the general proposition is the Second Circuit's decision in Russo v. Kirby, 453 F.2d 548 (2d Cir. 1971). In that case, striking workers were receiving welfare assistance. During the pendency of the strike, state officials altered their interpretation of state law, concluding that strikers were ineligible for welfare assistance. As a result, the striking workers previously receiving welfare funds were cut off. Rejecting a claimed violation

At the micro allocation level, the application of due process principles will depend much on the specific provisions of legislation and administrative regulation. In a program of entitlements, where a beneficiary is entitled to benefits upon meeting certain specified criteria, due process safeguards will continue to apply under *Goldberg v. Kelly*. Any similar medical benefits program that conferred benefits as a matter of right upon a beneficiary who satisfied certain objectively determinable criteria would also create a protected entitlement under due process. The "present enjoyment" limitation, however, could be construed to distinguish between an applicant for such a program and an actual beneficiary facing a benefits cutoff.[118]

An interesting problem arises when general eligibility criteria are established but no individual claims of entitlement are part of a program. An example of this is the Hill-Burton uncompensated care obligation. The Hill-Burton program requires a hospital receiving federal construction funds to deliver a reasonable amount of unreimbursed care to indigents in the community. The services mandated by Hill-Burton, however, do not entirely fulfill the prospective need. No "certain right to free services" was provided by the statute or the regulations "to any particular individual."[119] Consequently, no objective set of facts can automatically mandate that a particular individual receive Hill-Burton care at any given hospital, even if he satisfies the eligibility requirements with respect to income. As the Sixth Circuit held in *Newsom*, there is only the possibility of a benefit, not a claim of entitlement to it. Hospitals receiving Hill-Burton funds have the discretion to choose how to allocate their free care obligation,[120] and this discretion precludes the existence of an individual entitlement protected by the due process clause.[121]

The Seventh Circuit reached a contrary conclusion on the Hill-Burton entitlement issue in *Davis v. Ball Memorial Hospital Associa-*

---

of procedural due process, the court held that no constitutional rights of due process were implicated when state law or regulations were altered to affect the eligibility of a broad category of recipients. *Accord,* Connecticut State Dep't of Pub. Welfare v. HEW, 448 F.2d 209 (2d Cir. 1971).

Thus, where the government faces the need to diminish the resources allocated to an existing medical program, it can broadly rewrite eligibility standards or reduce benefit levels at the legislative or administrative levels without having to provide individual hearings as a matter of due process.

118. For further elaboration on some of these points, see subpart III(A)(4) *infra*.

119. *Newsom*, 653 F.2d at 1121.

120. Prior to 1979, federal regulations did not constrain hospitals' discretion in the method of allocating their unreimbursed care. *See id.* at 1119-21; Note, *supra* note 69, at 1504. The federal regulations adopted in 1979 require Hill-Burton hospitals to establish written criteria for allocating their unreimbursed care to indigent patients. 42 C.F.R. §§ 124.506-.507 (1981). "Depending upon whether the written eligibility criteria are subjective or objective, these written criteria may establish some claim of entitlement . . . ." *Newsom*, 653 F.2d at 1122.

121. From 1972 to 1979, federal Hill-Burton regulations provided for three methods of satisfying a hospital's unreimbursed care obligation. The hospital could furnish uncompensated services amounting annually to 3% of operating costs or 10% of all federal financial assistance received under the Hill-Burton program, whichever was less. In the alternative, a Hill-Burton hospital could satisfy its uncompensated care obligation by adopting an open-door policy, admitting all qualified indigents and providing them with uncompensated services. 42 C.F.R. § 53.111(d) (1977). The open-door option was eliminated under the 1979 Hill-Burton regulations. 42 C.F.R. § 124.503(a) (1981); *see* 44 Fed. Reg. 29,384 (1979). Consequently, a Hill-Burton hospital will no longer satisfy its uncompensated care obligation solely by agreeing to supply all needed services to qualified indigent patients. Hospitals that adopt the open-door approach, voluntarily supplying all needed services to all eligible indigent patients, may risk creating an entitlement.

tion[122] because of its belief that most applicants who meet the eligibility standards will in fact receive services. Therefore, the applicants would typically be required only to demonstrate their income, an objectively determinable factual inquiry. The Seventh Circuit assumed that the factual effect of the Hill-Burton program precluded any possible need to allocate scarce medical resources among an income-eligible population.

The Sixth Circuit in *Newsom* was unable to accept the factual assumption made by the Seventh Circuit. Instead, it saw the problem in light of the stipulation of the parties that the unsatisfied need for medical services was greater than either the resources of the hospital or its obligations under the Hill-Burton program.[123] Therefore, a mere showing of income eligibility would not be tantamount to establishing a right to treatment for any individual member of an indigent class. The Sixth Circuit concluded that due process principles were inapplicable because no claim of entitlement existed. To the extent that the Seventh Circuit found an entitlement in *Davis* by relying on the fact that most eligible Hill-Burton applicants would be approved for free care, the decision is in conflict with the Supreme Court's subsequent opinion in *Dumschat*, which held that the statistical probability of a successful application for a commutation did not create an entitlement to the commutation.[124]

Newsom's perceptive analysis seems to indicate that government can allocate funds to private parties on behalf of a broad group. Even if the subsequent actions of the private party can be attributed to the government, no claim of entitlement for any individual will exist and no due process requirements will apply.

3. Due Process Procedures.—If the threshold requirements of the fifth and fourteenth amendments—state action and deprivation of a protected interest—are met, then procedural due process guarantees will apply to a particular situation. Determining what process is constitutionally required depends upon an ad hoc balancing process. The Supreme Court has often noted that due process protection assures basic fairness in the adjudication process and is not a " 'technical conception with a fixed content unrelated to time, place and circumstances.' "[125] The requirements of due process have extended from the elaborate, trial-like procedures mandated in *Goldberg*[126] for welfare terminations to an informal give-and-take in school disciplinary proceedings.[127] In some situations, the Court has found no need for any hearing whatsoever when after-the-fact remedies satisfy the requirements of due process.[128]

122. 640 F.2d 30 (7th Cir. 1980).

123. *See Newsom*, 653 F.2d at 1121.

124. 101 S. Ct. at 2464. *See also* Jago v. Van Curen, 102 S. Ct. 31 (1981) (per curiam).

125. Little v. Streater, 452 U.S. 1, 5 (1981) (quoting Joint Anti-Fascist Refugee Comm. v. McGrath, 341 U.S. 123, 162 (1951) (Frankfurter, J., concurring)).

126. *See* 397 U.S. at 266-71.

127. *See* Goss v. Lopez, 419 U.S. 565, 584 (1975). *See also* Board of Curators v. Horowitz, 435 U.S. 78 (1978) (hearings not required for academic dismissals).

128. *See* Parratt v. Taylor, 451 U.S. 527 (1981) (state prison warden's negligent loss of an inmate's property does not deny due process when adequate state tort remedy is available); Ingra-

The Court established the general balancing approach in this area in *Mathews v. Eldridge*,[129] which involved the termination of disability benefits under Social Security:

> [I]dentification of the specific dictates of due process generally requires consideration of three distinct factors: First, the private interest that will be affected by the official action; second, the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional or substitute procedural safeguards; and finally, the Government's interest, including the function involved and the fiscal and administrative burdens that the additional or substitute procedural requirement would entail.[130]

The balancing formulation in *Mathews* means that each situation will require independent, ad hoc analysis. A few general comments are appropriate, however.

One procedural requirement often sought in a due process case is the adoption of formal standards to constrain the discretion of the decisionmaking body. Those seeking the adoption and publication of formal standards believe that such rulemaking serves to increase accountability and restrain governmental arbitrariness. A countervailing concern is the problem of legislative or administrative foresight and lack of flexibility. A genuine tension exists between rules and their perceived certainty, on the one hand, and individualized ad hoc decisionmaking and its resultant ambiguity on the other. The flexibility, possibly increased precision, and opportunity for the exercise of compassion in individualized decisionmaking counterbalance the benefits of the objectivity and limits on arbitrariness in formal standards.

The imposition of a requirement to articulate and publish allocative criteria creates problems for governmental decisions to ration scarce public dollars in the medical services field. Although some circuit courts in the 1960s held that public agencies had an obligation to adopt tenant selection criteria for public housing[131] and standards for the allocation of liquor licenses,[132] it is unclear whether these cases can survive the entitlements approach that has controlled due process adjudication since the decision in *Roth* in 1972.[133] We face a "chicken-and-egg" analytical problem. If no entitlement exists, then no due process applies. Yet it is difficult to see how an entitlement, as discussed in the last subsection, can exist unless objective allocative criteria are already in existence. General eligibility for a benefit is insufficient to warrant designation as an entitlement. Consequently, due process should not

---

ham v. Wright, 430 U.S. 651 (1977) (hearings need not precede corporal punishment of public school students when existing common-law remedies satisfy due process requirements).

129. 424 U.S. 319 (1976).

130. *Id.* at 335. In *Mathews* the Court declined to apply to the disability context the *Goldberg* requirement for a hearing prior to the termination of welfare benefits; a subsequent opportunity for a hearing was sufficient under due process when disability benefits were at stake.

131. *See* Holmes v. New York City Hous. Auth., 398 F.2d 262 (2d Cir. 1968).

132. *See* Hornsby v. Allen, 326 F.2d 605 (5th Cir. 1964).

133. Under the entitlements approach, the existence of an entitlement is a threshold inquiry. Absent such an entitlement, procedural due process does not apply. Consequently, there would be no analytical justification for requiring the articulation of standards unless some preexisting entitlement were found. *See generally* Bishop v. Wood, 426 U.S. 341 (1976).

require the articulation of allocative criteria because, in the absence of such criteria, no entitlement exists for due process purposes.

Of course, as a matter of policy, government administrators could choose to establish and publish allocative criteria for medical care. One likely outcome of such action would be the creation of an entitlement that would grant specified individuals the rudimentary due process rights of notice and a hearing. While the desirability of reducing arbitrariness is undeniable, the articulation of clearly formulated criteria poses substantial problems. Pragmatically, formulating standards is difficult because of the variety of problems and nuances that inevitably arise in the medical services context. More importantly, there is the specter of government openly renouncing the widely held belief in the sanctity of human life and publicly acknowledging that society is unwilling to expend the funds necessary to preserve human life and health. In short, as the kidney dialysis experience demonstrates,[134] the formulation of explicit priorities can have a debilitating effect on long-held symbolic values, especially since "hard-nosed" decisions often affect clearly identifiable individuals.

Proponents of the articulation of standards for allocating medical resources often maintain that political accountability and visibility are antidotes to both unfairness and potential or submerged discrimination[135] against "those groups within society who are traditionally the objects of neglect and mistreatment."[136] But unless nonsubstantive devices such as a lottery are used, some overtly inegalitarian values will of necessity appear in official allocative criteria. The acceptance of this inegalitarian choice will be made more difficult because of the very fact of its legitimation through the political process.[137] It would represent a politically sanctioned consensus that some illnesses will be favored over others or that some people will be more highly valued than others because of certain overtly articulated factors. This would constitute a significant breach of the societal myth of egalitarianism, which has been deservedly cherished as an abstract ideal.

The example of a wrongful death action well illustrates the difference in perception and impact between a clearly articulated, formally adopted set of allocative criteria and a less visible, ad hoc use of such criteria. A major component of proof in a wrongful death case is the determination of how much value to assign to the life of the decedent. One very important factor in this consideration is evidence on the lifetime earnings potential of the decedent. Obviously, if a defendant runs down a person of modest means and earning capacity, he will be re-

134. *See generally* Caplan, *Kidneys, Ethics, and Politics: Policy Lessons of the ESRD Experience*, 6 J. HEALTH POL., POL'Y, & L. 488 (1981).

135. *See, e.g.,* Note, *Due Process in the Allocation of Scarce Lifesaving Medical Resources*, 84 YALE L.J. 1734, 1747 n.61 (1975).

136. *See* J. KATZ & A. CAPRON, CATASTROPHIC DISEASES: WHO DECIDES WHAT? 3 (1975).

137. Of course, a lottery reflects a choice not to choose along substantive lines because no substantive criteria are adopted. For a discussion of the problem of selecting allocative criteria, see Sanders & Dukeminier, *Medical Advance and Legal Lag: Hemodialysis and Kidney Transplantation*, 15 U.C.L.A. L. REV. 357, 366-80 (1968); Note, *Scarce Medical Resources*, 69 COLUM. L. REV. 620, 639-66 (1969). *See generally* Furman v. Georgia, 408 U.S. 238, 314, 360-69 (1972) (Marshall, J., concurring).

quired to pay less for his mistake than if he had by accident run into a more affluent person. Nobody really thinks much about this situation or challenges it, even though the government's court system is implicated in a type of differential treatment that, in another context, might seem to be a form of elitism.

Imagine if the government were directly responsible for developing compensation formulas for wrongful deaths in a program of social insurance. Many would probably consider such perceived uneven treatment in the allocation process to be inequitable. After all, why should the amount of tort recovery depend upon whether the victim is well-to-do or a person of modest means? Indeed, if anything, one should compensate the family of the modest income-earner more because it probably needs the money more than the wealthy victim's family.

The wrongful death example raises questions about the "governmentalization" of allocative and ultimately life-valuing decisions by requiring the adoption of formal allocative standards. These criteria must involve choices among unpleasant alternatives, so that virtually any choice will have its own inevitable drawbacks. Yet the more visible and directly accountable the government is, the more nearly analogous to capital punishment decisions these issues become and the more they implicate the host of institutional and symbolic problems discussed in part II.

In these "tragic choice" situations, the use of what Professors Guido Calabresi and Philip Bobbitt call "a-responsible agencies" may serve a useful function.[138] Ad hoc, standardless decisions, like those of juries or hospital committees, for example, have an idiosyncratic character that disguises the weight being assigned to different variables and blurs the hierarchy of values chosen. The informal nature of these decisions and the lack of specifically articulated criteria mitigate the "devastating impact" of the "complex and hard choices" that are implicitly made.[139]

In conclusion, government would be imprudent to establish articulated rules or principles to govern micro decisions on rationing scarce and possibly lifesaving medical resources. A "tension between principle and expediency"[140] exists in a democratic system of government. "No good society can be unprincipled; and no viable society can be principle-ridden."[141] That would appear to be wise counsel with respect to the public articulation of explicit rationing criteria for medical care.

4. Benefit Reductions.—When governmental cutbacks in funding result in either stiffer eligibility standards or reduced benefits in an existing program, current recipients might have a sufficient entitlement to

138. *See* G. Calabresi & P. Bobbitt, Tragic Choices 57 (1978).
139. *See* J. Katz & A. Capron, *supra* note 136, at 3.
140. *See* A. Bickel, The Least Dangerous Branch 68 (1962).
141. *See* Bickel, *The Supreme Court, 1960 Term—Foreword: The Passive Virtues*, 75 Harv. L. Rev. 40, 49 (1961).

impose a duty of notice on the government. *Goldberg v. Kelly*[142] stressed the hardships caused to welfare beneficiaries by the termination of their benefits,[143] and either the termination or a significant reduction of benefits probably would qualify as a deprivation of an existing constitutionally protected entitlement.[144]

A strong argument against a finding of an entitlement exists if one focuses solely on the new eligibility or benefit standards. The government presumably can conserve resources by redefining eligibility standards or reducing benefit levels by administrative or legislative action. Under those circumstances, the particular details of the newly revised program will determine whether an entitlement exists. However, a beneficiary who has been receiving services and is adversely affected by a program change faces a deprivation of preexisting benefits that should trigger some form of due process protection. These beneficiaries, therefore, have a stake in being notified of any changes in eligibility standards or benefit levels. Because of their need to plan, they have a substantial interest in receiving timely notification of both the government's plans and the procedural opportunities available under the new guidelines.

Federal regulations in fact require "timely and adequate notice" to all affected recipients "[i]n cases of intended action to discontinue, terminate, suspend, or reduce assistance."[145] The regulations require that "timely" notice be mailed at least ten days before the effective date of the intended cutback.[146] In cases of automatic grant adjustments for classes of recipients, an "adequate" notice must include a statement of the intended action, the reasons for the intended action, the specific change in the law requiring such action, and a statement of the circumstances under which a hearing may be obtained and assistance continued.[147] These federal regulatory provisions were designed to implement the *Goldberg* decision[148] and serve the important interest of providing existing beneficiaries with notice of an impending action.

Even a threatened deprivation of existing benefits should constitute a sufficient interest to warrant procedural due process protection. A recipient's claim for notice is somewhat basic and relatively easy to meet. However, the question of whether an individual faced with a program cutback or termination must be given a hearing in order to comport with standards of due process depends upon whether the individual contests determinations of policy or determinations of fact. No hearing is required "when either State or Federal law require [*sic*] automatic grant adjustments for classes of recipients unless the reason for

142. 397 U.S. 254 (1970).

143. *Id.* at 264.

144. *See, e.g.,* Merriweather v. Burson, 325 F. Supp. 709, 710 (N.D. Ga. 1970), *remanded*, 439 F.2d 1092 (5th Cir. 1971).

145. 45 C.F.R. § 205.10(a)(4) (1980).

146. *Id.* § 205.10(a)(4)(i)(A).

147. *Id.* § 205.10(a)(4)(iii).

148. *See generally* Cooper, *Goldberg's Forgotten Footnote: Is There a Due Process Right to a Hearing Prior to the Termination of Welfare Benefits When the Only Issue Raised Is a Question of Law?*, 64 MINN. L. REV. 1107 (1980).

an individual appeal is incorrect grant computation."[149] An agency can deny or dismiss a hearing request "where the sole issue is one of State or Federal law requiring automatic grant adjustments for classes of recipients."[150] Consequently, under federal regulations applicable to Medicaid, an existing beneficiary faced with a macro policy change is entitled to a hearing only if he claims that the state erroneously computed the amount of or eligibility for a grant.[151]

This dichotomy between issues of policy and issues of fact arguably stems from the decision in *Goldberg* itself. The Supreme Court emphasized that pretermination hearings were important in cases, such as *Goldberg*, "where recipients have challenged proposed terminations as resting on incorrect or misleading factual premises or on misapplication of rules or policies to the facts of particular cases."[152] The Court reserved judgment on the due process requirements where no factual issues existed "or where the application of the rule of law is not intertwined with factual issues."[153] As a subsequent district court opinion noted, *Goldberg* seemed designed to reduce the opportunity for error where adjudicative facts concerning an individual's eligibility were in dispute. However,

> where a reduction or termination is not thus grounded on particular facts relating to an individual recipient . . . there is no need for an evidentiary hearing. Thus, where across-the-board cuts in funding necessitate wholesale reductions in benefits or . . . result in "automatic" reductions or terminations, it would be a useless expenditure of money to hold hearings at the request of any number of recipients opposed to reductions dictated by the state or federal legislature, rather than by the facts governing eligibility of particular recipients.[154]

The distinction between fact and policy determinations seems to rest on the belief that an aggrieved beneficiary should have a chance to demonstrate that he falls within a particular category established by law. However, there is no reason to afford a hearing to a beneficiary where the basis of a termination or reduction in his benefits is a political rationing choice adopted by the legislature or an administrative agency. If the fairness of that choice is to be challenged at all, it must be challenged substantively in a court of proper jurisdiction.

The distinction between facts and policy is usually analyzed in terms of what procedures due process requires. The distinction, however, also appears to have roots in the kind of interest that due process principles are designed to protect. In the case of macro eligibility standards or benefit levels, no individual claim to an entitlement to a new program exists, since state law defines the general entitlement at the outset. Only application of general macro principles to particular

149. 45 C.F.R. § 205.10(a)(5) (1980).

150. *Id.* § 205.10(a)(5)(v).

151. For a general review of these issues in the context of statewide welfare cutbacks, see Note, *Welfare, Due Process, and "Brutal Need": The Requirement of a Prior Hearing in State-Wide Benefit Reductions*, 34 VAND. L. REV. 173 (1981).

152. 397 U.S. at 268.

153. *Id.* at 268 n.15. *See also* Southern Ry. v. Virginia, 290 U.S. 190, 195 (1933).

154. Merriweather v. Burson, 325 F. Supp. 709, 711 (N.D. Ga. 1970).

factual circumstances warrants an evidentiary hearing. If an individual indisputably falls outside a redefined eligibility category, no entitlement protected by procedural due process exists. Thus, the distinction between facts and policy in the federal regulations is appropriate not only because of a determination of what procedures are fair, but also because of the nature of an entitlement as elaborated by post-*Roth* cases such as *Greenholtz*[155] and *Dumschat*.[156]

While the fact/policy distinction seems widely accepted,[157] its application is not free from ambiguity. Some courts have found the distinction to be unworkable,[158] while others have defined the range of factual questions very broadly.[159] Yet the principle is seen as valid if correctly applied.[160] This suggests that the government can effectuate a macro rationing policy. Existing recipients must receive notice of an impending action, but a computerized, fill-in-the-blank format is sufficient.[161] If factual matters are in dispute—for example, whether an individual falls within a specific category—an evidentiary hearing may be necessary. The issues, however, will be confined to the determination of eligibility or the computation of a grant.

The most difficult question concerns the requirements of due process when the new eligibility standards or benefit levels are linked to a discretionary formulation—for example, "in the public interest." Assume that a current recipient of public medical care is informed that he will no longer receive public funds because local decisionmakers conclude that he has less of a claim than others on public money under the "public interest" standard. If the rationing standard were objectively determinable—for example, if nobody over eighty years of age qualified—a hearing on the issue would be proper. The "public interest" standard seems sufficiently discretionary under *Greenholtz* to raise questions about the need for a hearing. The nature of the entitlement in such cases will be an important factor in determining whether a right to a hearing exists.

The limited factual scope of a hearing, the lack of a requirement of a hearing when only matters of policy are at stake, and the possible circumvention of a hearing requirement with broad, discretionary rationing standards suggest that considerable institutional flexibility exists to combat the thorny political and symbolic issues involved in rationing medical care.[162]

---

155. *See* text accompanying notes 94-95 & 100-03 *supra*.

156. *See* text accompanying notes 104-06 *supra*.

157. *See, e.g.*, Budnicki v. Beal, 450 F. Supp. 546, 555-56 (E.D. Pa. 1978). *But see* Mothers' & Children's Rights Org. v. Sterrett, 467 F.2d 797, 800 (7th Cir. 1972).

158. *See, e.g.*, Yee-Litt v. Richardson, 353 F. Supp. 996, 1000 (N.D. Cal. 1973), *aff'd mem. sub nom.* Carleson v. Yee-Litt, 412 U.S. 924 (1973).

159. *See, e.g.*, Viverito v. Smith, 474 F. Supp. 1122 (S.D.N.Y. 1979).

160. *See* Budnicki v. Beal, 450 F. Supp. 546, 556 (E.D. Pa. 1978).

161. *See* Turner v. Walsh, 435 F. Supp. 707, 713 (W.D. Mo. 1977), *aff'd mem.*, 574 F.2d 456 (8th Cir. 1978).

162. In cases such as kidney dialysis, where a cutback could be literally a matter of life or death, a poignant analogy to capital punishment could undoubtedly be made. *See* subpart II(E)(2) *supra*. For a skeptical evaluation of the Supreme Court's willingness to develop the life interest concept under procedural due process, see *Constitutional Perspectives*, *supra* note 31, at 247.

B. Rationing Public Funds: Substantive Due Process

The due process clause also operates as a substantive check on the exercise of governmental power. Although at one time substantive due process notions were widely invoked to limit governmental authority, the last forty-five years have taught us that the courts in general will not rigorously scrutinize governmental action under substantive due process. Courts will sustain legislation that serves a legitimate public interest if the means chosen to carry out the legislative purposes are rational. This statement of the "rational basis" standard of review does not fully convey the extent of the Supreme Court's deference to governmental action. In essence, substantive due process review is a mechanical process resulting almost always in the validation of the challenged governmental action.[163]

Two exceptions exist to this extraordinary level of deference. First, where the Supreme Court deems individual liberties to be fundamental, it requires the government to show that its legislation is necessary to achieve a compelling state interest.[164] Thus, the Court has called "well settled" the proposition that "if a law 'impinges upon a fundamental right explicitly or implicitly secured by the Constitution [it] is presumptively unconstitutional.' "[165] The second exception occurs in the rare cases in which the Supreme Court has recognized an affirmative obligation on the part of the government to provide access to specific services.[166] Some have argued that the government has an obligation to provide access to medical services,[167] but the Court has refused to recognize a constitutional foundation for such a right[168] outside of the prison context.[169] The abortion funding cases are most instructive and explicit in this regard.

In its landmark abortion decision in *Roe v. Wade*,[170] the Supreme Court held that the fourteenth amendment's concept of personal liberty protected a woman's fundamental interest in "privacy," which included her decision whether to terminate a pregnancy. Since the Court deemed the interest to be fundamental, the state could abrogate this interest only when necessary to promote a compelling state interest. As the Supreme Court subsequently reemphasized in *Maher v. Roe*,[171] "the right protects the woman from unduly burdensome interference with her freedom to decide whether to terminate her pregnancy. It im-

163. *See, e.g.,* North Dakota State Bd. of Pharmacy v. Snyder's Drug Stores, Inc., 414 U.S. 156 (1973); Ferguson v. Skrupa, 372 U.S. 726 (1963); Williamson v. Lee Optical, 348 U.S. 483 (1955).

164. *See* Roe v. Wade, 410 U.S. 113 (1973).

165. Harris v. McRae, 448 U.S. 297, 312 (1980) (quoting City of Mobile v. Bolden, 446 U.S. 55, 76 (1980) (plurality opinion)).

166. *Compare* United States v. Kras, 409 U.S. 434 (1973), *with* Boddie v. Connecticut, 401 U.S. 371 (1971). *See generally Constitutional Perspectives*, *supra* note 31, at 256-70.

167. *E.g.,* Kennedy, *The Responsibilities of the Government: A Legislative Viewpoint*, in IMPLICATIONS OF GUARANTEEING MEDICAL CARE 103, 108-09 (J. Perpich ed. 1975). For the view that the right to health care does not imply a right to equal access to medical treatment, see Fried, *supra* note 26, at 30-32.

168. *See* Harris v. McRae, 448 U.S. 297 (1980).

169. *See* Estelle v. Gamble, 429 U.S. 97 (1976).

170. 410 U.S. 113 (1973). The nature of the interest recognized in *Roe v. Wade* was an individual's freedom to make certain kinds of basic decisions without excessive governmental compulsion.

171. 432 U.S. 464 (1977).

plies no limitation on the authority of a State to make a value judgment favoring childbirth over abortion, and to implement that judgment by the allocation of public funds."[172]

Maher involved a state decision to withdraw funding for nontherapeutic abortions in its Medicaid program. The Court noted very explicitly that the state had no constitutional duty "to pay any of the medical expenses of indigents."[173] The case turned only on whether some putative right to an abortion changed that general principle. The Court found that the state's refusal to pay for certain abortions "place[d] no obstacles . . . in the pregnant woman's path to an abortion."[174] The woman's indigency, for which the state is not constitutionally responsible, posed the dilemma for her. Absent a governmental duty to provide medical services, its refusal to pay for the abortion created no new governmental restriction on the indigent's access to an abortion.

Since *Maher* involved nontherapeutic abortions, some believed that the Supreme Court would treat medically necessary abortions in a different way. The Court in *Harris v. McRae*[175] declined to adopt that distinction. It acknowledged a woman's interest in protecting her own health, but said that "it simply does not follow that a woman's freedom of choice carries with it a constitutional entitlement to the financial resources to avail herself of the full range of protected choices."[176] While a woman has a right to be free from governmental coercion regarding her decision to terminate a pregnancy, she does not have a claim on the public fisc to support her choice. Freedom from governmental intrusion in a private realm does not automatically establish "an entitlement to such funds as may be necessary to realize all the advantages of that freedom."[177] Government may not impose unnecessary obstacles in the path of a woman's abortion decision, but it has no affirmative obligation to eliminate preexisting obstacles such as poverty, which the government did not cause.[178]

The abortion funding decisions in *Maher* and *McRae* give the government broad powers in the distribution of its medical care resources. The two decisions draw an important distinction between an individual's right to be free from governmental interference, which may well receive close constitutional protection, and a claimed constitutional right to have government funds allocated to pay for private choices. Both *Maher* and *McRae* rejected the latter claim.

In addition to raising the alleged infringement on the privacy interest recognized in *Roe v. Wade*, indigents could argue that government has an affirmative duty to provide access to medical care for those who are unable to afford to pay for care. The Supreme Court has held that in some cases indigent individuals may have a claim to gov-

172. *Id.* at 473-74.
173. *Id.* at 469.
174. *Id.* at 474.
175. 448 U.S. 297 (1980).
176. *Id.* at 316.
177. *Id.* at 318.
178. *Id.* at 316.

ernment-supported access to certain services. This doctrine originally arose in the criminal justice context,[179] but the Court has extended it to include a limited number of other governmental duties.[180] The Court has recognized only a narrow range of cases in which it will impose an affirmative funding obligation. These cases are characterized by the importance of the services involved and their "constitutional connections";[181] the degree of state monopolization of access to them; and the element of total deprivation based on lack of funds.[182]

In *Maher*, the Court made it clear that the right of access cases would not be applied in the medical services context. Justice Powell for the majority tersely observed that the state had made no effort to monopolize the means for terminating pregnancies through abortions. It had merely refused to finance them. Unless the government's control over medical resources becomes absolute, analogous to its regulation of the court system or the criminal justice process, the Court seems unlikely to conclude that individuals have an entitlement to governmentally provided medical services. Consequently, government apparently is free to allocate its scarce medical resources in accordance with its own sense of priorities.

Ironically, the Supreme Court has only imposed an affirmative duty on government to provide medical services in one area—for prisoners. In *Estelle v. Gamble*,[183] a prisoner alleged that inadequate treatment of a back injury constituted cruel and unusual punishment in violation of the eighth amendment. Although the Court ruled against the prisoner, Justice Marshall distinguished the prisoner's claim, which was based on poor quality treatment and bad medical judgment, from an allegation of "deliberate indifference to a prisoner's serious illness" and his "serious medical needs," which would constitute cruel and unusual punishment under the eighth amendment.[184] The Court thus found that government has an "obligation to provide medical care for those whom it is punishing by incarceration,"[185] since an "inmate must rely on prison authorities to treat his medical needs."[186]

The irony of *Gamble* rests on the fact that an indigent in need of medical treatment becomes constitutionally entitled to it only if he is incarcerated. *Gamble* illustrates how difficult it is for government to deny medical treatment after the establishment of a governmental responsibility and an accompanying relationship of dependence, but, as a practical matter, the right to treatment recognized in *Gamble* will have

179. Douglas v. California, 372 U.S. 353 (1963); Griffin v. Illinois, 351 U.S. 12 (1956).

180. In Boddie v. Connecticut, 401 U.S. 371 (1971), the state had monopolized the procedures for legally dissolving a marriage. No alternative method for obtaining a divorce existed, and "given the basic position of the marriage relationship in . . . society's hierarchy of values," the Court found it a denial of due process for a state to deny access to its divorce court solely because of a person's indigency. *Id.* at 374. The Court in Little v. Streater, 452 U.S. 1, 16 n.12 (1981), noted that no forum other than a paternity hearing existed in which a putative father could demonstrate that a child was not his. Since the blood test under Connecticut's evidentiary rules was extremely important in assuring the fairness of the adjudicatory process, the Court held that the state had an obligation to provide access to a blood test. *Id.* at 14-17.

181. United States v. Kras, 409 U.S. 434, 444-45 (1973).

182. San Antonio Indep. School Dist. v. Rodriguez, 411 U.S. 1, 20-25 (1973).

183. 429 U.S. 97 (1976).

184. *Id.* at 104-05.

185. *Id.* at 103.

186. *Id.*

little impact outside the criminal justice system. In *Ingraham v. Wright*,[187] which involved the disciplinary paddling of a public school pupil, the Supreme Court held that the cruel and unusual punishment doctrine did not apply outside the criminal law area. Justice Powell expressly distinguished *Gamble* on the ground that the "prisoner and the schoolchild stand in wholly different circumstances, separated by the harsh facts of criminal conviction and incarceration."[188] Thus, the eighth amendment right to medical care does not limit governmental medical rationing outside the criminal justice system.

C. Rationing Public Funds: Equal Protection

The equal protection clause mandates that government treat similarly situated people in a like manner. When, in allocating scarce public medical care resources, government classifies a certain group as eligible for a given benefit, thereby implicitly excluding others, members of the disadvantaged group can require the state to articulate a "legitimately defensible difference"[189] to justify the disadvantageous distinction. The essential equal protection inquiry is the determination of the content of the state's burden of justification.

Equal protection doctrine imposes a heavy burden on states only when a classification impinges on a fundamental interest or applies a basis of classification that is constitutionally disfavored. In traditional equal protection analysis, the courts have afforded a presumption of validity to governmental classifications. The state must have a legitimate purpose, and the means used to achieve the purpose must be rationally related to it. However, the range of legitimate interests is broad, and the Court will not examine the weight of the interest. The burden rests with a plaintiff to show that a classification is "wholly irrelevant to the achievement of the State's objective."[190] The legislative rationale does not even have to be supported by empirical evidence; the state need only make a reasonable, logical case for its classification, even if its legislative judgment turns out to have no basis in fact.[191] Consequently, unless a plaintiff can find a reason for the application of a stricter standard of review, the Court will almost certainly sustain a challenged legislative classification.[192]

In the area of "economics and social welfare,"[193] the Court has been unwilling to become involved in the "intractable economic, social, and even philosophical problems presented by public welfare assistance programs."[194] The Court has refused to "second-guess state officials charged with the difficult responsibility of allocating limited

187. 430 U.S. 651 (1977).

188. *Id.* at 669.

189. *See* Ely, *Legislative and Administrative Motivation in Constitutional Law*, 79 YALE L.J. 1205, 1224 (1970).

190. McGowan v. Maryland, 366 U.S. 420, 425-26 (1961).

191. *See* Minnesota v. Clover Leaf Creamery Co., 449 U.S. 456 (1981).

192. In the last 40 years, the Supreme Court has applied this traditional test only once to strike down state legislation, and the Court subsequently overruled that case. New Orleans v. Dukes, 427 U.S. 297 (1976) (overruling Morey v. Doud, 354 U.S. 457 (1957)).

193. Dandridge v. Williams, 397 U.S. 471, 485 (1970).

194. *Id.* at 487.

public welfare funds among the myriad of potential recipients."[195] The Court has declined to label as fundamental such matters as education,[196] welfare,[197] and housing.[198] In the abortion funding cases, the Court declined to hold medical care to be a fundamental interest.[199] The Court, therefore, has subjected these social welfare programs to minimal scrutiny. The fundamental interests branch of the equal protection doctrine does not hold out much hope to a plaintiff challenging a medical resources rationing classification.

The disfavored classification theory holds more promise as a limitation on overtly egregious systems of classification.[200] Classifications deemed suspect, such as race,[201] must be justified as necessary to promote a compelling state interest.[202] In this mode of analysis, the Court examines the strength of the governmental interest involved, not merely its legitimacy. In addition, the Court carefully scrutinizes the means adopted by the state to achieve its goal and examines the availability of alternatives that avoid the use of suspect classifications. Finally, the Court does not apply a presumption of validity to suspect classifications, so the state faces a more difficult burden of justifying the classification. Where a racial classification has been implemented, at least to the disadvantage of a minority,[203] the state has a virtually impossible obstacle to overcome.[204] Consequently, a race-based classification, even for social welfare allocations, would almost certainly be held invalid.

Although the Supreme Court continues to maintain that it applies a two-tier level of scrutiny, rational basis and compelling interest,[205] the Court has also adopted what appears to be an intermediate standard of review in the case of disfavored classifications that the Court is unwilling to label suspect.[206] The clearest example of this intermediate level comes in the area of gender-based classifications. At one time four of the Court's nine Justices were willing to declare gender a suspect category.[207] Subsequently, the Court has backed away from this position.[208] Nevertheless, the Court's scrutiny in gender cases is clearly stricter than the traditional rational basis review. *Craig v. Boren*[209] es-

195. *Id.*

196. *See* San Antonio Indep. School Dist. v. Rodriguez, 411 U.S. 1 (1973).

197. *See* Dandridge v. Williams, 397 U.S. 471 (1970).

198. *See* Lindsey v. Normet, 405 U.S. 56 (1972).

199. *See* Harris v. McRae, 448 U.S. 297, 318 (1980); Maher v. Roe, 432 U.S. 464, 469 (1977).

200. *See, e.g.,* Loving v. Virginia, 388 U.S. 1 (1967).

201. *See, e.g.,* Loving v. Virginia, 388 U.S. 1 (1967); McLaughlin v. Florida, 379 U.S. 184 (1964); Bolling v. Sharpe, 347 U.S. 497 (1954).

202. *See* Graham v. Richardson, 403 U.S. 365, 375-76 (1971).

203. The use of racial criteria for so-called benign purposes may receive less strict scrutiny. *See* Fullilove v. Klutznick, 448 U.S. 448 (1980); Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265 (1978); United Jewish Orgs. v. Carey, 430 U.S. 144 (1977).

204. *See* Dunn v. Blumstein, 405 U.S. 330, 363-64 (1972) (Burger, C.J., dissenting).

205. *See* Massachusetts Bd. of Retirement v. Murgia, 427 U.S. 307 (1976).

206. *See, e.g.,* Wengler v. Druggists Mut. Ins. Co., 446 U.S. 142 (1980). *See generally* Gunther, *The Supreme Court, 1971 Term—Foreword: In Search of Evolving Doctrine on a Changing Court: A Model for a Newer Equal Protection*, 86 HARV. L. REV. 1 (1972); *see also* Massachusetts Bd. of Retirement v. Murgia, 427 U.S. 307, 317 (1976) (Marshall, J., dissenting).

207. *See* Frontiero v. Richardson, 411 U.S. 677, 682 (1973) (opinion of Brennan, J., joined by Douglas, Marshall, & White, JJ.).

208. *See, e.g.,* Rostker v. Goldberg, 101 S. Ct. 2646 (1981); Michael M. v. Sonoma County Superior Court, 450 U.S. 464, 468-69 (1981).

209. 429 U.S. 190 (1976).

tablished the controlling standard in gender cases: "To withstand constitutional challenge, . . . classifications by gender must serve important governmental objectives and must be substantially related to achievement of those objectives."[210] This test suggests that the Court will examine the strength of a state's interest, requiring it to be more than merely legitimate but less than compelling. The scrutiny of means under this standard affords less discretion for a state than the rational relationship approach, but more flexibility than the compelling interest formula.

The gender cases indicate that legislative classifications that rely on gender will undergo serious and not merely pro forma judicial scrutiny. Although the Court may be somewhat deferential to legitimately benign classifications toward women,[211] it will not hesitate to invalidate gender classifications based on or perpetuating sexual stereotypes. A medical rationing decision overtly based on gender would have to be justified under the *Craig v. Boren* standard, so the state would need important reasons to justify its use of a gender classification.

Equal protection analysis does not end with the recognition that the courts give greater constitutional scrutiny to suspect and disfavored classifications. Because overt racial classifications are rare, and such gender classifications are becoming less frequent, equal protection analysis often requires detailed inquiry to determine whether a classification is based on suspect or disfavored criteria. The Supreme Court has held that proof of both racial and gender classifications must include evidence of discriminatory intent. In *Washington v. Davis*,[212] the Court established the rule that proof of discriminatory intent is necessary to demonstrate a race-based classification. Discriminatory purpose, not disparate impact, is the core of the constitutional violation.[213] The Court in *Personnel Administrator v. Feeney*[214] applied the same standard of proof to gender-based discrimination. *Feeney* involved a challenge to Massachusetts' absolute lifetime preference for veterans in public employment. Even though plaintiffs showed that the preference had a vastly disproportionate impact on women, the Court declined to find a gender-based classification. The Court held that a law must be "designed" to prefer one sex over the other to qualify as a gender-based classification.[215] Plaintiffs must show that "the decisionmaker . . . selected or reaffirmed a particular course of action at least in part 'because of,' not merely 'in spite of,' its adverse effects upon an identifiable group."[216]

---

210. *Id.* at 197.

211. *See, e.g.,* Kahn v. Shevin, 416 U.S. 351 (1974).

212. 426 U.S. 229 (1976).

213. *Davis* was a constitutional challenge to the validity of a qualifying test administered to applicants for positions as police officers. Plaintiffs showed that four times as many blacks as whites failed the test. In reversing the court of appeals, the Supreme Court noted that equal protection is aimed at preventing racially discriminatory official conduct. *Id.* at 237.

214. 442 U.S. 256 (1979).

215. *Id.* at 273. The policy in *Feeney* distinguished between veterans and nonveterans, not between men and women. The Court acknowledged that the law's adverse effect on women was foreseeable but held that a finding of discriminatory purpose requires more than proof of "awareness of consequences." *Id.* at 279.

216. *Id.*

The discriminatory intent standard for proving racial and sexual discrimination has important implications for governmental allocation decisions. For example, a decision to cut back a research program on sickle cell anemia, which almost exclusively afflicts blacks, would have a disproportionate racial impact. Yet, under the approach in *Washington v. Davis*, the evidence of racially disproportionate impact, standing alone, would be insufficient to warrant a finding of a racial classification. The Court would evaluate the neutral rationale offered by governmental officials in determining the actual purpose of the decision. If the record revealed a justifiable and plausible nonracial basis for the decision, the Court would probably uphold it in the absence of evidence of discriminatory animus.[217] Thus, the government could justify a decision to terminate sickle cell anemia research by showing that scarce research resources would likely save more lives if allocated to other, higher priorities. This would be the case in spite of the disproportionate effect on blacks.

Suppose, however, that the rationale for the termination of the sickle cell research program were the following: Medical research is funded as an investment in preserving economically productive human lives; acting in rational pursuit of that objective, government analysts have concluded that since, statistically, blacks as a group have poorer economic prospects than whites, a research program oriented toward blacks is not cost effective. Such a rationale would trigger the strictest judicial review. While economic productivity might be a permissible criterion for allocating scarce research funds, using race as a proxy for economic productivity is precisely the kind of under- and over-inclusive, blunt method of classification that is constitutionally forbidden when the basis of distinction is suspect. If the goal of economic productivity is to be successfully pursued, more precise means have to be identified and used.[218]

---

217. *See* Village of Arlington Heights v. Metropolitan Hous. Dev. Corp., 429 U.S. 252 (1977).

218. A lively debate exists concerning the construction of the nondiscrimination provision of Title VI of the Civil Rights Act of 1964. 42 U.S.C. §§ 2000d to 2000e-17 (1976 & Supp. IV 1980). Section 601 provides as follows: "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance." *Id.* § 2000d. In Lau v. Nichols, 414 U.S. 563 (1974), the Supreme Court relied on federal guidelines in holding that the effects standard is the proper test for measuring discrimination under § 601. In Regents of the Univ. of Cal. v. Bakke, 438 U.S. 265, 352 (1978) (Brennan, White, Marshall, & Blackmun, JJ., concurring in the judgment in part and dissenting in part), the ongoing vitality of the *Lau* effects approach was questioned.

The meaning of § 601 is important in the consideration of hospital closings. If courts adopt the disproportionate impact theory, it could have far-reaching consequences for rationing programs designed to contain costs. A disproportionate impact standard might very well insulate hospitals servicing primarily minority populations from closure, even if they were losing vast sums of money. For a thorough discussion of this issue, see Bryan v. Koch, 492 F. Supp. 212 (S.D.N.Y.), *aff'd*, 627 F.2d 612 (2d Cir. 1980).

A majority of the circuits construe Title VI under the discriminatory intent standard, but a minority still apply a disparate impact standard. *Compare* Castaneda v. Pickard, 648 F.2d 989, 1000 (5th Cir. 1981), *and* Cannon v. University of Chicago, 648 F.2d 1104, 1109 (7th Cir.), *cert. denied*, 102 S. Ct. 981 (1981), *and* Guardians Ass'n v. Civil Serv. Comm'n, 633 F.2d 232, 270 (2d Cir. 1980) (Kelleher, J., concurring), *cert. granted*, 102 S. Ct. 997 (1982), *and id.* at 274 (Coffrin, J., concurring), *with* NAACP v. Medical Center, Inc., 657 F.2d 1322 (3d Cir. 1981) (en banc), *and* Guadalupe Org., Inc. v. Tempe Elementary School Dist. No. 3, 587 F.2d 1022, 1029 n.6 (9th Cir. 1978). On the general problem of hospital closings, see *Redefining Government's Role*, *supra* note 14, at 886-88; Note, *The Prima Facie Case and Remedies in Title VI Hospital Relocation Cases*, 65 CORNELL L. REV. 689 (1980).

D. Restricting the Use of Private Funds

To this point, this Article has focused on governmental decisions that restrict the allocation of public funds. As noted, government's power to establish criteria of allocation is very broad, unless the government intentionally disadvantages an identifiable group through the use of suspect or disfavored classifications. This subsection considers the ability of government to limit the expenditure of private funds. The issues to be discussed arise in three contexts: governmental restrictions on provider investment decisions through such programs as certificate-of-need (CON); governmental restrictions on the access of practitioners to the medical marketplace; and governmental limitations on private expenditures for medical services.

1. Provider Investment and Revenue Restrictions.—Spurred by federal health planning legislation, virtually all states have enacted CON laws.[219] Under federal guidelines, state agencies must review and approve the addition by a health care facility of a health service if it is either "associated with" a capital expenditure of any amount or if it "entails annual operating costs" of at least $75,000. In addition, a CON agency must review a facility's decision to terminate a health service if that termination is "associated" with a capital expenditure in any amount.[220]

Under the Nixon Administration's Economic Stabilization Program,[221] the federal government controlled wages and prices, restricted reimbursement costs of third-party payors, and limited the annual growth of hospital revenues.[222] A number of states operate prospective reimbursement programs, which establish ceilings on the amounts hospitals can receive from participating third-party payors.[223] A governmental agency sets or approves hospital charges, third-party payment rates, and total hospital revenue. Hospitals are "at risk" for cost overruns, so they have an incentive to be concerned with the financial implications of changes in quantity, quality, style of care, scope of services, and efficiency.

These regulatory programs are in widespread operation.[224] Except for a very unusual case, such as the decision by the North Carolina Supreme Court striking down its CON law on state constitutional

219. BUREAU OF HEALTH PLANNING, U.S. DEP'T OF HEALTH, EDUC. & WELFARE, DATES OF CERTIFICATE-OF-NEED ENACTMENT AND 1122 AGREEMENT (1980).

220. 42 C.F.R. § 123.404(a)(3) (1981).

221. *See generally* Ginsburg, *Inflation and the Economic Stabilization Program*, in HEALTH: A VICTIM OR CAUSE OF INFLATION? 31 (M. Zubkoff ed. 1976).

222. *See* Dowling, *Prospective Reimbursement of Hospitals*, 11 INQUIRY 163 (1974).

223. *See* Bauer, *Hospital Rate Setting—This Way to Salvation?*, in HOSPITAL COST CONTAINMENT 324 (M. Zubkoff, I. Raskin & R. Hanft eds. 1978); Hellinger, *An Empirical Analysis of Several Prospective Reimbursement Systems*, in HOSPITAL COST CONTAINMENT 370 (M. Zubkoff, I. Raskin & R. Hanft eds. 1978).

224. *See* F. SLOAN & B. STEINWALD, *supra* note 51, at 134-35 (1980); Biles, Schramm & Atkinson, *Hospital Cost Inflation Under State Rate-Setting Programs*, 303 NEW ENG. J. MED. 664 (1980); Coelen & Sullivan, *An Analysis of the Effects of Prospective Reimbursement Programs on Hospital Expenditures*, 2 HEALTH CARE FINANCING REV., Winter 1981, at 1; Cohen, *Experiences of a State Cost Control Commission*, in HOSPITAL COST CONTAINMENT 401 (M. Zubkoff, I. Raskin & R. Hanft eds. 1978); *Redefining Government's Role*, *supra* note 14, at 876-79.

grounds,[225] these laws seem valid under traditional federal constitutional principles.[226] Judicial review in this area would probably be most restrained.[227] Absent gross arbitrariness in the administration of a program or an obvious policy of confiscatory rate setting, these programs seem to be legitimate governmental efforts at limiting private investment and revenue generation in the health industry.

Governmental power, however, may be somewhat circumscribed in the restriction of private research activity. Some commentators have suggested the use of restrictions on technological advances both as a means of controlling technology-related health expenditures[228] and as a way of halting such potentially troublesome activities as recombinant DNA research, in vitro fertilization, and the like. As the previous subsections indicated, a governmental decision to stop funding research in some of these areas would not be likely to face serious constitutional challenge. Moreover, if only an economic benefit or even a potential improvement in human health were at stake, such a restriction should survive constitutional scrutiny. However, because of the link between a scientist's interest in research and the freedom of inquiry associated with the first amendment, the government could not justify the imposition of any broad-scale scientific moratorium absent reasons of overwhelming and imminent national importance and the lack of any realistic alternatives to achieve the national objective.[229]

2. Practice Restrictions.—Government can limit the accessibility of the marketplace to persons who do not meet certain basic licensure standards.[230] It has the power, similarly, to prescribe rules of practice whose violation will result in the loss of a license to practice. Moreover, government is able to expand or contract the ability of allied health professionals to engage in different forms of practice. Using

225. *In re* Certificate of Need for Aston Park Hosp., Inc., 282 N.C. 542, 193 S.E.2d 729 (1973) (finding due process deprivation from state law denying hospital right to operate except upon issuance of certificate of need); *see* Havighurst, *supra* note 51, at 1144 n.3.

226. Not even relatively activist courts have followed the North Carolina Supreme Court's decision striking down the North Carolina certificate-of-need law. *See* Goodin v. Oklahoma Welfare Comm'n, 436 F. Supp. 583 (W.D. Okla. 1977); Mount Royal Towers, Inc. v. Alabama Bd. of Health, 388 So. 2d 1209, 1215 (Ala. 1980); Merry Heart Nursing & Convalescent Home, Inc. v. Dougherty, 131 N.J. Super. 412, 330 A.2d 370 (1974).

227. Some judicial scrutiny under the antitrust laws may be appropriate. *See* National Gerimedical Hosp. v. Blue Cross, 101 S. Ct. 2415 (1981); Huron Valley Hosp., Inc. v. City of Pontiac, No. 79-1265 (6th Cir. Dec. 11, 1981). *See generally* Blumstein & Calvani, *State Action as a Shield and a Sword in a Medical Services Antitrust Context:* Parker v. Brown *in Constitutional Perspective*, 1978 DUKE L.J. 389; Miller, *Antitrust and Certificate of Need: Health Systems Agencies, the Planning Act, and Regulatory Capture*, 68 GEO. L.J. 873 (1980).

228. For consideration of the relationship between costs and technology, see Gruenberg, *The Failures of Success*, 55 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 3 (1977); Hiatt, *Protecting the Medical Commons: Who Is Responsible?*, 293 NEW ENG. J. MED. 235 (1975); Iglehart, *The Cost and Regulation of Medical Technology: Future Policy Directions*, 55 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 25 (1977); Hiatt, *Too Much Medical Technology?*, Wall St. J., June 24, 1976, at 16, col. 4. For a discussion of the constitutional limits on governmental restriction of research, see Robertson, *The Scientist's Right to Research: A Constitutional Analysis*, 51 S. CAL. L. REV. 1203 (1977).

229. *See* Robertson, *supra* note 228.

230. *See generally* Cohen, *Professional Licensure, Organizational Behavior, and the Public Interest*, 51 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 73 (1973); Forgotson, Roemer & Newman, *Licensure of Physicians*, 1967 WASH. U.L.Q. 249. The ability of government to restrict a person's entry into the practice of a profession is not unlimited, however. *See* Special Project, *Admission to the Bar: A Constitutional Analysis*, 34 VAND. L. REV. 655 (1981).

public funds, it can influence the number of new practitioners who will emerge from medical schools and post-graduate training programs. And it is in a position to limit the number of new practitioners who enter the market in any given year.

In this way, government can constrain the way in which the medical marketplace serves individuals. Thus, even if a patient wants a midwife to deliver her baby, the state has the power to prohibit the practice of midwifery. The Constitution protects neither the right of the midwife to practice her profession nor the right of the patient to choose her provider.[231] Courts would probably uphold such restrictive state laws under the principles of substantive due process.

In addition, government can circumscribe the ability of practitioners to follow their own medical judgment in certain situations. Thus, it can outlaw particular medical procedures or medications, such as laetrile,[232] unless the restrictions have an unconstitutional impact upon patients' fundamental interests, as in contraception or abortion.[233] Until the government approves a drug, physicians may not prescribe it.[234] Presumably, the government may bar a physician from prescribing marijuana to a cancer patient, even if the doctor believes the drug is a legitimate means of medical treatment.[235] This type of restriction on the practice of a profession is valid under current substantive due process principles.[236] Indeed, the government can require the doctor to provide information about the prescription of narcotics to patients if the state deems such information to be an important mechanism for enforcing drug abuse laws.[237]

Although some of the Court's statements in the abortion cases suggest some type of constitutional protection for a physician's right to practice,[238] the Court apparently will not afford that right significant protection independent of the patient's interest in her life and health. Physicians unsuccessfully raised the right-to-practice claim in challenging the validity of the Professional Standards Review Organization (PSRO) law.[239] When the government pays, it can set standards for

231. *See generally* England v. Louisiana State Bd. of Medical Examiners, 246 F. Supp. 993, 995 (E.D. La. 1965), *aff'd per curiam*, 384 U.S. 885 (1966); Kentucky Ass'n of Chiropractors, Inc. v. Jefferson County Medical Soc'y, 549 S.W.2d 817 (Ky. 1977).

232. *See generally* United States v. Rutherford, 442 U.S. 544 (1979).

233. *See, e.g.,* Planned Parenthood v. Danforth, 428 U.S. 52, 75-79 (1976) (holding unconstitutional state prohibition on the use of saline amniocentesis as an abortion technique). *See also* Carey v. Population Servs. Int'l, 431 U.S. 678 (1977).

234. *See* United States v. Rutherford, 442 U.S. 544 (1979).

235. In Hartz v. Bensinger, 461 F. Supp. 431 (E.D. Pa. 1978), a cancer patient requested three marijuana cigarettes daily as doctor-supervised medication to be taken while the patient underwent chemotherapy. Although Judge Pollak denied the motion for a temporary restraining order, he indicated that the plaintiff's claim warranted "respectful attention." *Id.* at 433. The acknowledgment that the patient already had access to marijuana to get the sought after symptomatic relief apparently influenced Judge Pollak's decision. The court focused entirely on the plaintiff's interest in receiving marijuana to relieve the pain caused by his condition and its treatment. Consequently, any potential right that might emerge in this field apparently would be analogous to the contraception or abortion situations, in which courts view governmental restrictions on procedures or therapies from the perspective of the patient, not from the perspective of a doctor's right to practice.

236. *See generally* People v. Privitera, 23 Cal. 3d 697, 591 P.2d 919, 153 Cal. Rptr. 431 (1979).

237. *See* Whalen v. Roe, 429 U.S. 589 (1977).

238. *E.g.,* Doe v. Bolton, 410 U.S. 179, 198-200 (1973).

239. Association of Am. Physicians & Surgeons v. Weinberger, 395 F. Supp. 125, 131-35 (N.D.

practice and compensation. Where the courts have accorded weight to the right-to-practice notion, it has been linked with a claim of patients' rights.[240]

A facet of the traditional right to practice involves the physician's right to choose patients. Under common-law principles, the doctor-patient relationship is contractual. This is not an iron-clad rule—doctors in an emergency room may hold themselves out to patients in need and in many jurisdictions have a duty to treat emergency patients. In non-emergency settings, however, physicians typically have the freedom to reject patients. Of course, ambiguous situations arise concerning whether a doctor-patient relationship has been formed. Also, doctors cannot terminate such a relationship without assuring adequate continuity of care.[241]

Here the constitutional question is whether a doctor's traditional freedom to choose patients enjoys constitutional protection. That is, what would happen if government were to require that physicians accept all patients, including Medicaid and Medicare patients, on a non-discriminatory basis? It is unlikely that a court would sustain their claim that a right to practice, asserted against a patient's interest in receiving treatment, would constitutionally invalidate the requirement. Provided the rate of compensation were not patently confiscatory, government, as a matter of constitutional law, could probably enforce a requirement that physicians treat Medicaid and Medicare patients at specified fee levels. In short, government has considerable regulatory authority over the private practice of any business or profession. The major limitation, as reflected in the abortion cases, is the ability of the professional to represent, as a spokesman in court, the fundamental interests of the patient or client. Consideration of those interests will conclude this section.

3. Limits on Patients' Medical Expenditures.—Governmental restrictions on the ability of patients to spend their own funds on medical treatment can arise in several contexts. We have already seen that government can restrict the right of patients to choose providers in the marketplace. Similarly, patients do not have carte blanche at the pharmacy; government can restrict access to medication by requiring prescriptions or by banning substances altogether.

The government would have more difficulty justifying a limitation on the ability of private individuals to spend their own funds on medi-

---

Ill.), *aff'd mem.*, 423 U.S. 975 (1975). *See also* American Medical Ass'n v. Weinberger, 395 F. Supp. 515 (N.D. Ill.), *aff'd*, 522 F.2d 921 (7th Cir. 1975).

240. *See* American Medical Ass'n v. Weinberger, 395 F. Supp. 515, 524 (N.D. Ill.), *aff'd*, 522 F.2d 921 (7th Cir. 1975).

241. *See, e.g.,* Lyons v. Grether, 218 Va. 630, 239 S.E.2d 103 (1977). Section 5 of the Principles of Medical Ethics of the American Medical Association states:

> A physician may choose whom he will serve. In an emergency, however, he should render service to the best of his ability. Having undertaken the care of a patient, he may not neglect him; and unless he has been discharged he may discontinue his services only after giving adequate notice. He should not solicit patients.

AMERICAN MEDICAL ASSOCIATION, OPINIONS AND REPORTS OF THE JUDICIAL COUNCIL 4-5 (1977). *See generally* J. KING, JR., THE LAW OF MEDICAL MALPRACTICE 20-27 (1977).

cal care.[242] In a highly regulatory environment, in which regulators tightly constrained increases in capacity, one can envision a pressure for excessive utilization. If private use of medical facilities by some created bottlenecks or inaccessibility for others, outcome equity might dictate patterns of equal utilization irrespective of wealth. These concerns—supply constraint and outcome equity—could lead to consideration of governmental restrictions on the private consumption of medical care.

Government expenditure of public funds to support private medical services is quite a different thing from governmental prohibition of an individual's private pursuit of health care. The Supreme Court drew this distinction in the abortion funding cases. One theme of the abortion cases has been the importance of a woman's ability to pursue her own health unfettered by coercive governmental regulation. To the extent that the expenditure of private funds may be necessary in the pursuit of an individual's health, public limits on private medical spending arguably impinge upon a fundamental liberty interest analogous to the right of privacy recognized in the abortion cases. The private pursuit of individual health is distinguishable from business and professional judgments about investment in medical resources. The courts might find the individual's interest in the pursuit of health to be sufficiently fundamental to warrant a stricter standard of scrutiny.[243]

Other than the abortion cases, there is not much in the way of direct precedent because no such governmental restriction has been implemented. However, two cases dealing with education provide a helpful analogy. In the post-World War I period, a number of states sought to "Americanize" their immigrant populations by prohibiting instruction of youngsters in foreign languages[244] and requiring students to attend public schools. In the landmark decision of *Pierce v. Society of Sisters*,[245] the Court invalidated an Oregon law that required children to attend public schools, ruling that parents have a fundamental liberty interest in choosing a private rather than a public school for their children. Although the analogy between education and health is imprecise, the decisions in *Meyer v. Nebraska*[246] and *Pierce* suggest that the government faces limits in confining individuals' participation to or in a publicly controlled and supported system of social service delivery.[247] If the Court found the private interest in pursuing health to be fundamental, then the Court would have to examine carefully the state's coercive limitation on spending to determine how much of a deprivation

242. This would include a patient's interest in receiving symptomatic treatment even if the course of treatment is not within the medical mainstream. *See* Rogers v. State Bd. of Medical Examiners, 371 So. 2d 1037 (Fla. Dist. Ct. App. 1977).

243. *See* note 235 *supra*.

244. *See* Meyer v. Nebraska, 262 U.S. 390 (1923) (law prohibiting the teaching of German a violation of substantive due process).

245. 268 U.S. 510 (1925).

246. 262 U.S. 390 (1923).

247. One is tempted to draw an analogy to the campaign finance area. The Supreme Court has held that the expenditure of money is protected as a necessary part of political expression. *See* Buckley v. Valeo, 424 U.S. 1 (1976).

was involved, the magnitude of the governmental interest, and the availability of alternative policies to achieve the governmental purpose.

To the extent that governmental restriction on private funds took the form of a change in favorable tax treatment, the element of coercion, as that term was used in the abortion and abortion funding cases, would be absent. Governmental withdrawal or restructuring of a tax benefit to achieve an alteration in private spending patterns certainly would not pose the problem of a flat prohibition or cap on private health spending.

IV. Legal Implications of Private Rationing Decisions

A. Informed Consent

The relationship between patient and physician is consensual,[248] and the patient is formally the ultimate decisionmaker. In order to preserve the patient's right of self-determination, the law requires the patient to consent to the performance of any medical procedures. Except in unusual or emergency situations, a physician cannot legally treat a patient until the patient has given his consent. For that consent to be valid, it must be "informed."[249]

Two general approaches exist for defining the nature and scope of a physician's duty to disclose information when seeking a patient's consent to a medical procedure. The traditional and still majority view only requires a physician to disclose the information that a reasonable, prudent physician would disclose in similar circumstances.[250] This traditional approach relies on a professional frame of reference for determining the necessary degree of disclosure. The rationale is that given the complex nature of medical diagnosis and treatment decisionmaking, the propriety of risk disclosure is best left to the professional judgment of physicians,[251] whose decisions are to be made in accordance with professional norms.

A recent but still minority approach focuses on what information a reasonable patient in a given set of circumstances would want to know in order to make a decision. The District of Columbia Court of Appeals in *Canterbury v. Spence*[252] became the first court to apply this patient-oriented standard. Under the *Canterbury* approach, the focus of the disclosure requirement shifts from the professional's frame of reference to that of the patient. Instead of examining what is custom-

248. *See generally* A. HOLDER, MEDICAL MALPRACTICE LAW 1-18 (1975).

249. The doctrine of informed consent is founded on the principle, articulated by Judge Cardozo, that "[e]very human being of adult years and sound mind has a right to determine what shall be done with his own body." Schloendorff v. Society of N.Y. Hosp., 211 N.Y. 125, 129, 105 N.E. 92, 93 (1914), *overruled on other grounds*, Bing v. Thunig, 2 N.Y.2d 656, 143 N.E.2d 3, 163 N.Y.S.2d 3 (1957); *see* J. LUDLAM, INFORMED CONSENT (1978); Waltz & Scheuneman, *Informed Consent to Therapy*, 64 NW. U.L. REV. 628 (1970); Note, *Restructuring Informed Consent: Legal Therapy for the Doctor-Patient Relationship*, 79 YALE L.J. 1533 (1970).

250. *See, e.g.*, Natanson v. Kline, 186 Kan. 393, 409, 350 P.2d 1093, 1106 (1960).

251. *See* Karp v. Cooley, 349 F. Supp. 827 (S.D. Tex. 1972), *aff'd*, 493 F.2d 408 (5th Cir. 1974).

252. 464 F.2d 772 (D.C. Cir. 1972). *See also* Cobbs v. Grant, 8 Cal. 3d 221, 502 P.2d 1, 104 Cal. Rptr. 505 (1972); Wilkinson v. Vesey, 110 R.I. 606, 295 A.2d 676 (1972). For a general discussion of the functions of informed consent, see Capron, *Informed Consent in Catastrophic Disease Research and Treatment*, 123 U. PA. L. REV. 340, 364-76 (1974).

ary within the medical profession, the inquiry centers on what reasonable patients would want to know. The patients determine the extent of their participation in risk assessment decisions. The professional, and often paternalistic, judgments of doctors do not control malpractice liability. As a practical matter, proof is easier since lay witnesses can testify on the disclosure issue.

A governmental decision to ration public dollars may mean that public third-party financing will be unavailable for a certain procedure, but that may not be determinative of the physician's duty to disclose.[253] The fact that the government assigns a low priority to a certain procedure for collective purchases does not necessarily indicate that a particular individual would assign the same value to it. A *Canterbury* jurisdiction would require a physician to disclose any information that a reasonable patient would want to know about alternative courses of treatment, regardless of the availability of third-party reimbursement. Therefore, in a *Canterbury* jurisdiction, governmental rationing decisions would probably not be decisive in determining the scope of the duty to disclose.

In a traditional jurisdiction, a professional custom might well emerge that the governmental rationing decision reflects an appropriate standard for decisionmaking. If the government would not pay for coronary artery bypass surgery without the presence of certain surgical indications, the norm in the medical community might not require disclosure of the surgical option until the government's required surgical indications had been satisfied. Thus, if a reasonable, prudent cardiologist would not have informed a patient of the surgical possibility, then no duty to disclose would exist in a non-*Canterbury* jurisdiction. Consequently, governmental decisions could influence the scope of the duty to disclose in non-*Canterbury* jurisdictions more easily than in *Canterbury* jurisdictions.

Any consideration of the decisions of private physicians to ration medical care independent of governmental policy must focus on the nature of the physician's duty to the patient. A physician who has entered into a doctor-patient relationship is obliged to act as a fiduciary in the patient's interest. In a non-*Canterbury* jurisdiction, the nature and scope of disclosure would depend on a professional frame of reference. However, the standard probably must take as its starting point the patient's medical interest. A decision by doctors, as a group, that disclosure of nonserious and unusual complications was unwarranted would face little public opposition. It would be quite something else, even in a non-*Canterbury* jurisdiction, to accept a disclosure principle based not on a medical assessment of the patient's interest but on an assessment of society's best interest in the overall allocation of resources. Such an approach would represent a basic challenge to the principle of patient self-determination that underlies the informed consent doctrine. It would also threaten in a fundamental way the integrity of the doctor-patient relationship by undermining the fiduciary

253. *Cf.* Havighurst & Blumstein, *supra* note 37, at 55-58 (suggesting this approach for care disapproved for federal funding by a PSRO).

responsibility of the physician to allow the patient to exercise at least formal control over the decisionmaking process.[254]

Although physicians in non-*Canterbury* jurisdictions have more flexibility in determining the nature and scope of their duty to disclose, there may still be limitations on their nondisclosure of diagnostic risks or alternative treatment opportunities if nondisclosure is based on a sense of overall (macro) allocation priorities, rather than on a sense of what is appropriate for a patient's making an individualized (micro) medical decision. A doctor might be obliged to inform a patient that an extra day in intensive care would be medically desirable even though a hospital staff committee had decided that the intensive care space could be better utilized or a third-party carrier would not reimburse the patient for the extra care rendered. At least the patient would be able to make the ultimate decision about his medical care—do nothing, pay the amount not covered by the third-party carrier, convince hospital officials of the importance of additional intensive care, or even transfer to another hospital if possible. The formation of a doctor-patient relationship should trigger the disclosure duty. Consequently, absent a therapeutic rationale and the accompanying therapeutic privilege,[255] a physician's duty to disclose should extend to prospective treatment options as well as to decisions about the withdrawal of care.

B. Malpractice Standards: Physicians and Institutions

The law of medical malpractice is designed to achieve the twin objectives of quality assurance and victim compensation.[256] The quality assurance goal is promoted by reliance on input and process standards of medical care delivery.[257] Thus, the proof in a malpractice action will focus on the training of a professional, the facilities of an institution, and the process-oriented aspects of quality assurance, such as medical recordkeeping, use and interpretation of indicated laboratory tests, and prescription of customary or accepted treatment regimens.[258] Given the relative deemphasis on outcome measures of quality in malpractice litigation, a provider who seeks to ration medical care resources—for example, by ordering fewer or less costly tests, by providing a smaller margin of safety in terms of facilities or equipment availability, or by allowing less highly credentialed personnel to perform certain procedures—runs the risk of increased exposure to malpractice liability in case of a medical maloccurrence.

A malpractice inquiry has two components. The first necessitates the establishment of a standard of care. The second examines a physi-

254. Doctors do not always act solely in the interest of their patients. For example, insurance companies assign physicians to examine persons to determine the extent of a disability. For a discussion of the function of physicians in controlling access to the social benefits associated with the sick role, see Stone, *Diagnosis and the Dole: The Function of Illness in American Distributive Politics*, 4 J. Pol., Pol'y, & L. 507 (1979).

255. *See generally* Meisel, *The "Exceptions" to the Informed Consent Doctrine: Striking a Balance Between Competing Values in Medical Decisionmaking*, 1979 Wis. L. Rev. 413.

256. *See* King, *In Search of a Standard of Care for the Medical Profession: The "Accepted Practice" Formula*, 28 Vand. L. Rev. 1213, 1227-34 (1975).

257. *Id.* at 1234-41.

258. For further elaboration of the difference between the customary and the accepted practice concepts, see *id.* at 1234-44.

cian's application of the standard of care in a given situation.[259] Rationing decisions may have an impact on the first issue, but it is doubtful that rationing would substantially affect the second.

The establishment of a proper standard of care in a malpractice lawsuit is based on expert testimony about what a reasonable, prudent physician would have done in similar circumstances. As with the doctrine of informed consent, the prevailing rule focuses on a professional standard. Only in the rarest of circumstances will a court find the controlling professional standard to be inadequate.[260]

Physicians are liable if they breach a duty to a patient, but the nature of the duty depends upon how one defines the controlling standard of care. In general, the appropriate standard is derived by referring to what other similarly situated professionals would do in similar circumstances. A professionally developed standard does not in all cases immunize professional behavior that complies with the standard—"there are precautions so imperative that even their universal disregard will not excuse their omission"[261]—but very few cases have rejected the conclusiveness of the professional standard in medical malpractice.[262]

Although the professional standard controls virtually all medical situations, some controversy exists about its articulation. Traditionally, expert testimony would establish the standard of care by describing the customary behavior of physicians in a given set of circumstances. The customary standard is established, therefore, by developing empirical evidence of actual patterns of medical practice. Under that approach, no normative judgment enters into the analysis to determine whether the existing practice standard is medically desirable.

An alternative approach focuses on "accepted" practice,[263] a concept that suggests a normative element—that a "customary" standard of practice in an area may not reflect medical practices approved by the profession at large. The "accepted practice" formulation places emphasis not on what would be done by someone else but on what is "expected of a reasonably competent practitioner . . . ."[264]

The existence of a "respectable minority" rule may offset the significance of the distinction between the customary and accepted practice standards. This rule acknowledges both that medical practice is not monolithic and that opinions on effective diagnostic and treatment regimens differ. To give flexibility to the exercise of clinical judgment, the "respectable minority" doctrine holds that "a physician does not incur liability merely by electing to pursue one of several recognized courses of treatment."[265] If respectable medical authority supports a certain course of conduct, a physician is allowed to follow that mode of

259. *See* J. King, Jr., *supra* note 241, at 37.

260. *See In re* Quinlan, 70 N.J. 10, 25-26, 44-47, 355 A.2d 647, 655, 666-67 (1976) (decision to remove a comatose patient from a respirator may conform to broader moral standards of the community at large rather than to professional medical standards); Helling v. Carey, 83 Wash. 2d 514, 519 P.2d 981 (1974).

261. The T.J. Hooper, 60 F.2d 737, 740 (2d Cir. 1932).

262. *But see* Helling v. Carey, 83 Wash. 2d 514, 519 P.2d 981 (1974).

263. *See generally* King, *supra* note 256, at 1236-44.

264. Blair v. Eblen, 461 S.W.2d 370, 373 (Ky. 1970).

265. Downer v. Veilleux, 322 A.2d 82, 87 (Me. 1974).

practice without breaching a duty of care to the patient. Of course, a doctor may have to disclose alternative treatment options to a patient in a *Canterbury* jurisdiction, but that is a question of informed consent.

Governmental rationing decisions that affect patterns of medical practice are likely to have an effect on the applicable standard of care, whether "customary" or "accepted." Under the "respectable minority" approach, if even a minority of practitioners decide to adjust their practice standards to conform with governmental rationing guidelines, then arguably a follower of that school would have a legitimate defense to a malpractice claim. The federal PSRO legislation contains a provision that immunizes from malpractice liability a physician who practices in conformity with the standard set forth by the PSRO.[266] The clear intent of that provision was to remove the "defensive medicine" excuse from practitioners who claim that fear of malpractice liability causes them to order more costly testing than ordinarily called for in their clinical judgment.[267] The PSRO provision does not immunize a practitioner from negligence in carrying out a set standard of care, but it does allow the PSRO norm to serve as the basis for the physician's own delivery of service. Being able to rely on the PSRO norm eliminates the need for utilization of unnecessary procedures as a precaution against malpractice liability.

The existence of the malpractice immunity provision of the PSRO law is an attractive vehicle for the government to use as a cost containment measure in any future rationing of medical care. The government can obviate the risk of malpractice without encouraging defensive medicine by having the relevant PSRO adopt the appropriate rationing standard. That standard would then supersede any conflicting state common-law standard. Moreover, even in cases where the PSRO immunity provision did not apply, the very existence of a PSRO standard in an area would be solid evidence to support the stature of that approach under the "respectable minority" rule.

Where private medical practitioners seek to engage in cost containment rationing on their own, they risk running afoul of either the customary or accepted practice rule. Malpractice law emphasizes input measures as a means of quality assurance. An avant garde practitioner who declines to order certain tests that other physicians routinely order runs the risk that a harmed patient can successfully sue for malpractice. The procedure orientation of malpractice law makes it difficult for a practitioner to conclude, on his own, that a certain treatment costs more than it is worth. Unless the physician can point to a respectable minority viewpoint in the medical literature or in general practice, he leaves himself open to a malpractice claim. It is this risk of liability that the PSRO immunity provision overcomes where it is applicable, primarily in the Medicare and Medicaid context.

266. 42 U.S.C. § 1320c-16(c) (1976). *See generally* Note, *Federally Imposed Self-Regulation of Medical Practice: A Critique of the Professional Standards Review Organization*, 42 GEO. WASH. L. REV. 822, 838-42 (1974); Comment, *PSRO: Malpractice Liability and the Impact of the Civil Immunity Clause*, 62 GEO. L.J. 1499 (1974).

267. For a discussion of the cost implications of the civil immunity provision, see Blumstein, *The Role of PSROs in Hospital Cost Containment*, in HOSPITAL COST CONTAINMENT 461, 473 (M. Zubkoff, I. Raskin & R. Hanft eds. 1978).

In addition to physicians, institutional health care providers are subject to malpractice liability in appropriate circumstances. For example, a hospital can be held vicariously liable for the negligence of its employees, both physicians and nonphysicians. Typically, hospitals are liable only for the negligence of physicians who are actual employees, not for the acts of doctors with staff privileges. Some courts have breached this distinction, however, under the "ostensible agency" doctrine, which applies if the hospital's actions reasonably lead a patient to conclude that a physician is an employee and the patient justifiably relies on the physician's services.[268]

A hospital also is liable to patients for its own negligence. In jurisdictions that follow the Illinois decision in *Darling v. Charleston Community Memorial Hospital*,[269] hospitals as corporate entities owe a duty of care to patients to provide, among other things, adequate and safe equipment, reasonable levels of staffing, and qualified personnel. Thus, a hospital must "exercise reasonable care in accordance with sound hospital practice to protect the health and safety of its patients."[270]

Apparently no requirement exists that all hospitals carry every piece of the latest available equipment.[271] However, a hospital may be obliged to have equipment adequate to perform a service it holds itself out as capable of performing. The jury in a malpractice trial ultimately decides the standard of reasonableness as a matter of fact on an ad hoc, case by case basis. Thus, where a hospital purports to limit access to a given resource, such as an intensive care unit, the reasonableness of the scope of the intensive care facility itself may be a legitimate bone of contention in a negligence action. It would be an appropriate matter of inquiry whether the decision to limit the availability of intensive care facilities was a reasonable, prudent exercise of the hospital's duty to its patients. Of course, the existence of CON laws would be a relevant part of the inquiry, since a hospital undoubtedly could use the denial of its request to expand as a defense. The interrelationship of CONs and institutional liability has yet to emerge in the cases,[272] however, and the federal CON legislation does not include a malpractice immunity provision like that of the PSRO legislation.

Finally, principles of institutional liability developed for hospitals seem applicable to HMOs, which are prepaid group practice plans,[273] despite their somewhat unique set of incentives. Because an HMO provides medical care to a group of patients, it must live within a fixed budget set by individual capitation payments. HMOs, therefore, have a powerful incentive to balance costs, risks, and outcomes in deciding

268. J. KING, JR., *supra* note 241, at 296.

269. 33 Ill. 2d 326, 211 N.E.2d 253 (1965).

270. J. KING, JR., *supra* note 241, at 294.

271. *Id.* at 309.

272. One very interesting area of interrelationship might be the antitrust laws, where CON legislation might confer on a provider market power that might not otherwise exist. For further development of this theme, see Blumstein & Calvani, *supra* note 227, at 431-37; *Redefining Government's Role*, *supra* note 14, at 920-24.

273. *See* Curran & Moseley, *The Malpractice Experience of Health Maintenance Organizations*, 70 NW. U.L. REV. 69 (1975).

what services to provide. The existence of a predetermined budget induces HMOs to evaluate carefully the effectiveness and efficiency of their procedures. HMOs have an incentive to look at outcomes for their entire enrolled population, rather than at specific procedures for a specific patient. Yet, traditional malpractice standards focus on input measures of quality at the level of the individual doctor-patient encounter.

Apparently, "[t]he malpractice standard of care applied to HMO health care services is the same as that governing medical services generally."[274] Since one virtue of HMOs is their nonconventional mode of treatment, application of the customary practice approach to HMOs may impose liability inappropriately. A major concern with applying the customary practice rule to HMOs is that the traditional standards are derived from insured fee-for-service medical practice, in which providers make medical judgments with little regard for costs.[275] Mechanical application of customary practice standards to HMOs could reduce their ability to adopt innovative styles of practice. Perhaps the respectable minority principle can offset this risk for HMOs, or perhaps courts can be persuaded to adopt an "HMO custom" approach to govern malpractice actions against HMOs. This would protect individuals and at the same time permit the development and emergence of different patterns of HMO practice.[276]

V. Conclusion

Health care institutions and the proponents of equitable access to medical services have combined to develop a health care system that is structurally biased toward the consumption of more and better medical services. The increasing costs attributable to health care, partly generated in response to the perceived problem of inadequate and inequitable access to medical care, have led to concern for cost containment efforts. Yet calls for centralized rationing, especially by the government, raise their own institutional problems of deeply held, conflicting viewpoints in a sensitive, ideologically charged, political environment. The heavy overlay of symbolism makes accommodation of the conflicting values particularly difficult.

From a policy perspective, the preferable course of action would be to decentralize decisionmaking in order to avoid invoking the symbolic issues that inevitably arise when the government is perceived as directly responsible for a person's life or health. Nevertheless, if some form of governmental policy is necessary, it should seek to distance the government from nitty gritty micro allocation matters to the greatest extent possible by developing macro allocation policies and leaving their implementation to nongovernmental units where appropriate.

---

274. Bovbjerg, *The Medical Malpractice Standard of Care: HMOs and Customary Practice*, 1975 DUKE L.J. 1375, 1386.

275. *Id.* at 1397.

276. *See id.* at 1408-14. For a discussion of the relationship between HMOs and PSROs, see Havighurst & Bovbjerg, *Professional Standards Review Organizations and Health Maintenance Organizations: Are They Compatible?*, 1975 UTAH L. REV. 375.

Rationing "Normal" Health Care: The Hidden Legal Issues

Rand E. Rosenblatt*

Does American law restrict the rationing of health care by government and private health care providers? In his article on this topic,[1] Professor James Blumstein argues that the Constitution imposes virtually no limits on such government rationing[2] decisions as reducing the coverage of medical assistance for the poor,[3] regulating hospital investment in expensive equipment,[4] and refusing to pay for abortions.[5] Although he concedes that government procedures for rationing health benefits in individual cases are theoretically subject to the constitutional restraints of procedural due process, he argues that recent cases allow cleverly drafted statutes to override these procedural rights.[6] Blumstein does believe, however, that malpractice law does or might deter providers, particularly Health Maintenance Organizations (HMOs), from rationing, i.e., from deviating from costly, inefficient fee-for-service practices.[7] He proposes that government delegate its rationing authority—a power he argues it cannot use effectively because of the symbolism surrounding life and death decisions—to private providers through such devices as partial civil immunity legislation and

* Associate Professor of Law, Rutgers University Law School (Camden). B.A. 1966, Harvard University; M.Sc. (Econ.) 1967, London School of Economics and Political Science; J.D. 1971, Yale University.

An earlier version of this Article was prepared for the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research [hereinafter referred to as the Commission]. The views represent those of the author, not those of the Commission.

1. Blumstein, *Rationing Medical Resources: A Constitutional, Legal, and Policy Analysis*, 59 TEXAS L. REV. 1345 (1981).

2. Although Blumstein does not clearly define rationing, he speaks of it both as a "method" of controlling health care costs and as "criteria for allocating scarce medical resources." *Id.* at 1346.

3. *Id.* at 1356 n.50, 1373-74.

4. *Id.* at 1356 n.51, 1386-88.

5. *Id.* at 1378-80.

6. *Id.* at 1370-73 & text accompanying notes 142-56.

7. *Id.* at 1395-99.

Published in 59 TEXAS L. REV. 1401 (1981). Copyright © 1981 by the Texas Law Review. Reprinted by permission.

discretionary grants.[8] He maintains that the dangers of provider discrimination in the rationing process will be minimized by civil rights legislation, tort law, and public scrutiny.[9]

This Article takes issue with Blumstein on two points. First, he ignores the complicated relationship between rationing and social class. Needed care is widely denied to low-income and minority patients, who are often financially excluded from private, nonprofit hospitals and frequently face substandard conditions in public hospitals.[10] The statutes, regulations, and malpractice standards that in theory guard against these problems in fact offer little protection.[11] Second, Blumstein fails to deal adequately with the unequal relationship between doctors, hospitals, and insurance companies on the one hand, and patients and nonphysician providers on the other. The powerful actors in the health care industry have historically defined and controlled rationing and allocation decisions, often to the detriment of patients from all social classes.[12] Because of the impact of social class and professional/industry power, one cannot assume that government delegation of rationing authority to private providers will result in fair or efficient rationing.[13]

---

8. *Id.* at 1397. *See also id.* at 1354, 1369.

9. *Id.* at 1363.

10. *See, e.g.,* Rosenblatt, *Dual Track Health Care—The Decline of the Medicaid Cure*, 44 U. CIN. L. REV. 643, 652, 653 & n.44 (1975) [hereinafter cited as *Dual Track Health Care*]; Rosenblatt, *Health Care Reform and Administrative Law: A Structural Approach*, 88 YALE L.J. 243, 270-71 & nn.105-06 (1978) (citing hospital's alleged denial of care to the seriously ill despite Hill-Burton Act commitments) [hereinafter cited as *Health Care Reform*]; Sparer, *Gordian Knots: The Situation of Health Care Advocacy for the Poor Today*, 15 CLEARINGHOUSE REV. 1, 10-11 (1981).

11. *See* text accompanying notes 61-88 *infra*.

12. *See Dual Track Health Care*, *supra* note 10, at 644-45 (nineteenth-century charitable discretion); *Health Care Reform*, *supra* note 10, at 287-88 (pre-1965 state and local welfare discretion); *id.* at 305-06 & n.294 (pre-1974 health planning efforts); Payton & Powsner, *Regulation Through the Looking Glass: Hospitals, Blue Cross, and Certificate-of-Need*, 79 MICH. L. REV. 203 (1980) (describing industry domination in development of regional planning and certificate of need legislation).

13. The bulk of Blumstein's article is devoted to describing current constitutional law and the implications for rationing health care. Rather than focus in depth on actual health care policies, he concentrates on defending the general proposition that government has great substantive and procedural latitude in allocating funds and in regulating investment, licensing, and treatment. *See* Blumstein, *supra* note 1, at 1356-92. As a description of where the current majority of the Supreme Court stands on constitutional issues, Blumstein's discussion is generally accurate. He does not, however, mention the major criticisms of the Court's decisions and reasoning. For example, Blumstein discusses with apparent approval the cases upholding state and federal refusal to pay for the abortions of poor women, *see* Blumstein, *supra* note 1, at 1378-80, without acknowledging how inconsistent these decisions appear to be with Roe v. Wade, 410 U.S. 113 (1973), and with other cases concerning the impact of funding decisions on constitutional rights. *See, e.g.,* Perry, *Why the Supreme Court Was Plainly Wrong in the Hyde Amendment Case: A Brief Comment on* Harris v. McRae, 32 STAN. L. REV. 1113 (1980). More fundamentally, the Court's decisions cited with approval by Blumstein have been criticized as systematically biased against the poor, minorities, and other politically weak groups. *See, e.g.,* L. TRIBE, AMERICAN CONSTITUTIONAL LAW 1127-1136 (1978); Clune, *The Supreme Court's Treatment of Wealth Discriminations Under the Fourteenth Amendment*, 1975 SUP. CT. REV. 289; Tushnet, . . . *And Only Wealth Will Buy You Justice—Some Notes on the Supreme Court, 1972 Term*, 1974 WIS. L. REV. 177; Note, *Cases That Shock the Conscience: Reflections on Criticism of the Burger Court*, 15 HARV. C.R.-C.L. L. REV. 713 (1980). Blumstein obviously does not share this critical perspective on the Court's work, but it deserves at least a response.

Blumstein's failure to discuss the realities of contemporary rationing of health care for the poor and the middle class reduces the value of his doctrinal discussion. What the Court would do if confronted squarely with some of the more egregious examples of withholding or restricting health care is not as clear as Blumstein implies; what the Court *should* do is even more debatable. For example, Blumstein asserts with great confidence that even "if a patient wants a midwife to deliver her baby, the state has the power to prohibit the practice of midwifery. The Constitution

The thesis of this Article is that there are actually two types of rationing that affect the distribution of health services: the traditional type still applied to the poor, and a new type said to be applicable to patients of all social classes. Traditional rationing is wholly unjustifiable because it has virtually no relationship to a population's actual needs or to priorities for meeting them.[14] The new rationing is in theory based on an analysis of actual health needs, but it contains both harmful and beneficial potential. American law could regulate both types of rationing in beneficial ways, but it has unfortunately failed to do so, with adverse consequences for both poor and middle class patients.

I. Rationing Health Care: Images and Realities

For many Americans, rationing conjures up images of ration books and the other regulatory devices of World War II. The distinctive features of wartime rationing were scarcity of supply and a national judgment that goods should be distributed according to collective standards of priority and fairness, rather than solely according to ability to pay. Thus everyone was entitled to purchase a fixed amount of basic goods—such as butter and meat—at regulated prices, while high-priority military functions received extra shares. Wartime regulation contrasts sharply with most people's picture of the peacetime market, in which consumption is perceived as depending upon consumer preferences and ability to pay and in which the concept of "rationing" has no place.[15]

Whatever the accuracy of the market image for most goods and services, it has never seemed very plausible when applied to medical and hospital care. For many health services, need is urgent, information scarce, and performance standards uncertain, thereby precluding the possibility of informed consumer choice.[16] In addition, much of the population cannot—and historically could not—afford to pay for health care out of personal savings and current income.[17]

The traditional response of American communities and the medical profession to economic inequality was not to provide lower income

---

protects neither the right of the midwife to practice her profession nor the right of the patient to choose her provider." Blumstein, *supra* note 1, at 1388 (citation omitted). In support of this position, Blumstein had cited two cases upholding restrictions on the licensing and practice of chiropractors. *See id.* at 1388 n.231 (citing England v. Louisiana State Bd. of Medical Examiners, 246 F. Supp. 993, 995 (E.D. La. 1965), *aff'd per curiam*, 384 U.S. 885 (1966); Kentucky Ass'n of Chiropractors, Inc. v. Jefferson County Medical Soc'y, 549 S.W.2d 817 (Ky. 1977)). But even were these cases persuasive precedent, a far stronger case could be made for the value and distinctiveness of midwifery, and for the fundamental nature of patient rights in childbirth decisions. *Cf.* Fitzgerald v. Porter Memorial Hosp., 523 F.2d 716, 723-24 (7th Cir. 1975) (Sprecher, J., dissenting) (hospital policy prohibiting presence of father at childbirth infringes patient's constitutional rights to procreative choice and to choice of physician and treatment). *See also* note 94 *infra* & accompanying text.

14. *See* text accompanying notes 18-20 & 30-33 *infra*.

15. *See* J. BLUM, V WAS FOR VICTORY 227 (1976).

16. *See, e.g.,* Rosenblatt, *Health Care, Markets, and Democratic Values*, 34 VAND. L. REV. 1067, 1079-82 (1981) [hereinafter cited as *Market Failures*].

17. *See, e.g.,* COMMITTEE ON THE COSTS OF MEDICAL CARE, MEDICAL CARE FOR THE AMERICAN PEOPLE 2-35 (1932 & reprint 1972); S. LAW, BLUE CROSS: WHAT WENT WRONG? 6-11 (2d ed. 1976) (describing the development of Blue Cross as a mechanism to cover medical costs); *Health Care Reform*, *supra* note 10, at 264-65.

patients with the means to purchase normal or "mainstream" services, but rather to establish separate hospitals, wards, and clinics for the poor.[18] The existence of this "dual track" system of care required pervasive rationing in two senses. First, lower income patients were usually excluded from the personalized medical care associated with traditional standards of quality.[19] Second, the separate institutions for the poor were badly underfunded, and scarce resources were rationed by lowering standards of care or by denying it altogether to those in need.[20]

During the past decade, rationing has come to have two other meanings for health policy. The more familiar meaning refers to the allocation of scarce, high technology treatment among the relatively large class of patients who would benefit from it.[21] The other new meaning of rationing provides the focus for this Article and for Blumstein's: it pertains to efforts to discourage the use of "normal" health care resources, such as diagnostic tests, surgery, drugs, and hospitals.

The impetus for the second type of new rationing lies in the changing structure of the health care industry. The expansion of insurance for hospital and most medical services, including government-financed programs for the aged (Medicare) and some of the poor (Medicaid), has lowered the financial barrier to increasingly complex and expensive care. Reimbursement policies have given physicians and hospitals a strong economic incentive to perform more services at higher prices. The medical profession's emphasis on specialization and technology as the hallmarks of "quality" has reinforced this trend. The result has been a rapid escalation of the costs of medical and hospital services, with little assurance that the hundreds of billions of dollars involved are being wisely spent.[22] In response to this situation, health care analysts have suggested various strategies for discouraging the use of health care services by most patients.[23] The sense of scarcity that previously applied only to the poor is thus now being extended—with very different goals and techniques—to patients with average and above-average incomes.

---

18. *See Dual Track Health Care*, *supra* note 10, at 644-45; Rosenberg, *Social Class and Medical Care in 19th Century America: The Rise and Fall of the Dispensary*, 29 J. HIST. MED. 32 (1974), *reprinted in* SICKNESS AND HEALTH IN AMERICA 157 (J. Leavitt & R. Numbers eds. 1978).

19. For a discussion of the historical patterns of medical care, see Rosenberg, *The Practice of Medicine in New York A Century Ago*, 41 BULL. HIST. MED. 223 (1967), *reprinted in* SICKNESS AND HEALTH IN AMERICA, *supra* note 18, at 55; Rosenberg, *supra* note 18. Raymond Duff and August Hollingshead, investigating a major teaching hospital in the Eastern United States in the 1960s, found this pattern to be in many ways unchanged, with patients from higher social classes receiving much greater personal attention and comprehensive medical care. *See* R. DUFF & A. HOLLINGSHEAD, SICKNESS AND SOCIETY 124-50, 371-72 (1968).

20. *See Dual Track Health Care*, *supra* note 10, at 644-45.

21. In defining the scope of inquiry of the papers on health care rationing, the Commission assigned this subject a lesser priority, *see* Letter from Mary Ann Baily, Staff Economist to the Commission, to Professor James Blumstein (June 4, 1981), at 3 (copy on file with the *Texas Law Review*) [hereinafter cited as Bailey letter], and Blumstein's article does not treat it extensively, *see generally* Blumstein, *supra* note 1, at 1386-92.

22. *See, e.g., Health Care Reform*, *supra* note 10, at 244-45 & nn.4-7; Roberts & Bogue, *The American Health Care System: Where Have All the Dollars Gone?*, 13 HARV. J. ON LEGIS. 635 (1976); Stevens, *High Medical Costs Under Attack As Drain on the Nation's Economy*, N.Y. Times, Mar. 28, 1982, § 1, at 1, col. 1.

23. *See, e.g., Market-Oriented Approaches to Achieving Health Policy Goals*, 34 VAND. L. REV. 849 (1981); Roberts & Bogue, *supra* note 22; Stevens, *supra* note 22.

The new rationing differs from the old in important ways. Under the old model, the allocation of resources to the poor was almost never based on an analysis of their actual needs; instead, governments and charitable donors decided how much they wished to give and provided care within those limits.[24] Moreover, the standards for allocating the scarce services were often unrelated to need, but were instead based on traditional concepts of who "belonged" in a geographic area, judgments of social worthiness, and exercises of unreviewable discretion.[25] In contrast, the new rationing theoretically begins with an analysis of the community's actual health needs and seeks to organize resources and financial incentives to meet those needs effectively and efficiently.[26] Proponents of both regulatory and market reforms justify their proposals not as ways of denying patients needed care, but rather as ways of expanding our collective and individual capacities to identify what care is truly needed.[27] To be sure, such ambitious claims have been subjected to strong criticism, both on the ground that the new techniques rest on flawed assumptions about health care and human behavior[28] and because they have been frequently "captured" by the dominant forces in the health care industry.[29] Without attempting to resolve such issues here, it is sufficient to note that the new rationing embodies substantially different goals than the old rationing and proposes new techniques for achieving them.

The development of a new concept of rationing has not, however, displaced or eliminated the rationing practices long applied to the poor. On the contrary, many of the developments that have fueled escalating costs and declining access for all patients have reinforced traditional patterns of exclusion and inequality. As hospital equipment and medical procedures become more costly, hospitals have found it increasingly difficult to provide charitable care, and patients without adequate insurance have been turned away.[30] The decline of the general practitioner and the rise of the specialist with more expensive training and expectations of high-income have further isolated low-income patients

24. *See, e.g., Dual Track Health Care*, *supra* note 10, at 644-45 & n.6; *Health Care Reform*, *supra* note 10, at 287-88.

25. *See Health Care Reform*, *supra* note 10, at 287-88 & n.194.

26. *See, e.g.,* V. FUCHS, WHO SHALL LIVE? 143-51 (1974).

27. *See, e.g., Health Care Reform*, *supra* note 10, at 304-05; *Market Failures*, *supra* note 16, at 1077-78.

28. *See Market Failures*, *supra* note 16, at 1067-69, 1078-1115.

29. *See, e.g., Health Care Reform*, *supra* note 10; Payton & Powsner, *supra* note 12.

30. The clearest legal recognition of this trend has occurred in requirements for federal tax exemption. Under a 1956 Revenue ruling, a hospital seeking tax-exempt status as a charitable institution under § 501 of the Internal Revenue Code had to be "operated to the extent of its financial ability for those not able to pay for the services rendered and not exclusively for those who are able and expected to pay." Rev. Rul. 56-185, 1956-1 C.B. 202, *quoted and discussed in* Eastern Ky. Welfare Rights Org. v. Shultz, 370 F. Supp. 325, 328 n.5 (D.D.C. 1973), *rev'd sub nom.* Eastern Ky. Welfare Rights Org. v. Simon, 506 F.2d 1278 (D.C. Cir. 1974), *vacated on other grounds*, 426 U.S. 26 (1976). This Internal Revenue Service ruling itself reflected a long-standing hospital tradition of providing charitable care to those unable to pay. *See, e.g., Eastern Ky. Welfare Rights Org.*, 370 F. Supp. at 331. In 1969, however, the hospital industry "actively endorsed and advocated" a bill that would have granted federal tax exemption to "any" hospital regardless of its service to indigents. *Id.* at 332 n.13. Although Congress refused to enact such an amendment, *id.*, the Service promulgated a new ruling granting tax-exempt status to a hospital that admitted persons unable to pay only to its emergency room, and "otherwise ordinarily limits admissions to those who can pay the cost of their hospitalization, either themselves or through

from mainstream medical care.[31] As the cost of medical services has risen, the state and federal governments have grown increasingly reluctant to pay the bills, even for the selected categories of recipients—the aged, the disabled, and low-income single-parent families—eligible for Medicare and Medicaid.[32] For the most part, the result has been a perpetuation of the traditional dual track system in which underfunded urban public hospitals and clinics, and a small number of private practitioners organized as "Medicaid mills," provide the bulk of services to the poor.[33]

II. Professor Blumstein and the Problem of the Vanishing Poor

Professor Blumstein largely ignores the continued existence and importance of traditional rationing practices for the poor.[34] The reader of his article is given no hint, for example, that patients who urgently need medical and hospital care are often denied it because they cannot afford to pay for it,[35] or because health care providers discourage them from seeking it on grounds of race and other "social characteristics." Blumstein does not mention hospital admissions policies requiring pa-

---

private health insurance [or public programs] . . . ." Rev. Rul. 69-545, 1969-2 C.B. 193. In addition, Revenue Ruling 69-545 explicitly eliminated the requirement of free or below-cost service established by the 1956 Ruling. A federal district court held that Revenue Ruling 69-545 was invalid because it was inconsistent with congressional intent regarding charitable exemptions for hospitals, 370 F. Supp. at 330-38, but the district court was reversed by the court of appeals, in part on the grounds that "soar[ing] hospital costs and expanded governmental programs had reduced both the feasibility and necessity of traditional charitable care. *See* 506 F.2d 1278, 1288 (D.C. Cir. 1974), *vacated on other grounds*, 426 U.S. 26 (1976). The Supreme Court vacated the court of appeals' judgment and dismissed the complaint on the ground that the plaintiffs lacked standing to challenge Revenue Ruling 69-545. *See* Simon v. Eastern Ky. Welfare Rights Org., 426 U.S. 26, 46 (1976). The result is that Revenue Ruling 69-545 remains in force.

For other examples of exclusion of patients without adequate insurance, see *Financial Triage in Nashville*, HEALTH PAC BULL., Nov.-Dec. 1981, at 16 (Baptist Hospital, largest voluntary private-nonprofit hospital in Nashville, Tenn., undertakes $45 million expansion program while successfully resisting proposal to require care for the poor and minorities; hospital now requires $1,000 advance deposit for admission). *See also* note 35 *infra*.

31. *See, e.g.,* Rogers, *The Challenge of Primary Care*, in DOING BETTER AND FEELING WORSE 81, 84-86 (J. Knowles ed. 1977).

32. *See, e.g., Health Care Reform*, *supra* note 10, at 292-95; Stevens, *supra* note 22.

33. *See, e.g., Dual Track Health Care*, *supra* note 10, at 652, 653 & n.44; *Health Care Reform*, *supra* note 10, at 244 nn.5-6.

34. In defining the topic for inquiry, the Commission implicitly recognized the coexistence of the traditional and new meanings of rationing. Rationing was preliminarily defined as "limiting the consumption of [beneficial treatment.]" Baily letter, *supra* note 21, at 1. The examples given by the Commission referred both to limitations imposed on all patients as part of insurance, hospital, or HMO policy (examples of new rationing) and to refusals by hospitals and physicians to treat certain patients at all (an example of traditional rationing). *Id.* at 1, 3. The Commission was interested in the general topic of how American law affected such broadly defined rationing, and especially in the argument, often advanced by physicians, that malpractice law prohibits any rationing that withholds beneficial treatment for reasons other than the patient's welfare. *Id.* at 2.

35. *See, e.g.,* Newsom v. Vanderbilt Univ., 453 F. Supp. 401 (M.D. Tenn. 1978), *aff'd in part, modified in part, and rev'd in part*, 653 F.2d 1100 (6th Cir. 1981). The district court in *Newsom* concluded, after examination of numerous hospital records, that

> applicants [for care] are led to believe, and realistically so, that they will be turned away if they are perceived to be unable to pay. Moreover, of the many people who have been denied admission for financial reasons, a significant number have had nowhere else to turn for the needed treatment.

Id. at 417. Professor Edward Sparer reports that the records of Vanderbilt University Hospital list 211 persons turned away for financial reasons during an eleven-month period. *See* Sparer, *supra* note 10, at 10. *See also Health Care Reform*, *supra* note 10, at 271 n.107 (quoting from plaintiff's post-trial brief in *Newsom*, which in turn quoted hospital records.)

tients to have insurance other than welfare,[36] pay cash deposits,[37] or establish a relationship with a "private physician."[38] Nor does he discuss hospital decisions to close outpatient clinics or relocate to the suburbs,[39] physician decisions not to accept Medicaid patients,[40] or the substandard, dangerous conditions in many major municipal hospitals.[41] All of these factors effectively deny needed care to the poor, sometimes with tragic consequences.[42]

The persistence of old-style rationing for the poor is apparent not only in explicit exclusionary rules, but also in the pattern of access to expensive care. For example, in 1981 an estimated 100,000 to 125,000 Americans underwent coronary artery bypass grafting (CABG), at an average cost of $10,000 to $25,000 per operation and an aggregate cost of at least two billion dollars.[43] While serious questions have been raised about whether CABG prolongs life or improves the quality of life for many patients,[44] it is widely regarded as a leading treatment for heart disease, "the ultimate weapon against premature infirmity or death."[45] The distribution of this controversial "benefit" among the population, however, is apparently inconsistent with projected health needs. Two major studies conducted in the 1970s found that ninety-seven percent of CABG patients were white.[46] "Yet according to the [National Center for Health Care Technology] . . . age-adjusted mortality statistics suggest that the prevalence of coronary artery disease is similar in blacks and whites. And, in fact, angina tends to disable

36. *See, e.g.,* Cook v. Ochsner Found. Hosp., 61 F.R.D. 354, 364 (E.D. La. 1972). *See also Health Care Reform*, *supra* note 10, at 270 nn.104-05.

37. *See Financial Triage in Nashville*, *supra* note 30; *Health Care Reform*, *supra* note 10, at 270 nn.104-05.

38. *See, e.g.,* Campbell v. Mincey, 413 F. Supp. 16 (N.D. Miss. 1975); Perry v. Greater S.E. Community Hosp. Found., No. 725-71 (D.D.C. June 28, 1972), *reprinted in* 2 HEALTH LAW PROJECT OF THE UNIVERSITY OF PENNSYLVANIA LAW SCHOOL, MATERIALS ON HEALTH LAW: THE HOSPITAL 71 (rev. ed. 1972) [hereinafter cited as THE HOSPITAL] (discussed in *Health Care Reform*, *supra* note 10, at 271); Cook v. Ochsner Found. Hosp., 61 F.R.D. 354, 364-65 (E.D. La. 1972).

39. *See, e.g.,* NAACP v. Wilmington Medical Center, Inc., 491 F. Supp. 290 (D. Del. 1980), *aff'd*, 657 F.2d 1322 (3d Cir. 1981) (en banc).

40. *See, e.g.,* Holahan, Hadley, Scanlon, Lee & Bluck, *Paying for Physician Services Under Medicare and Medicaid*, 57 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 183, 198-99 (1979) (finding that more than 50% of a sample of 5003 California physicians do not participate in Medicaid).

41. *See, e.g.,* Greater Washington, D.C. Area Council of Senior Citzens v. District of Columbia, 406 F. Supp. 768 (D.D.C. 1975); Boone v. Tate, Civ. No. 713 (C.P. Philadelphia County, Apr. 30, 1971), *reprinted in* THE HOSPITAL, *supra* note 38, at 440; *Dual Track Health Care*, *supra* note 10, at 653 n.44.

42. *See, e.g.,* Stanturf v. Sipes, 447 S.W.2d 558 (Mo. 1969) (refusal by nonprofit community hospital to treat indigent patient for frostbite led to amputation of both feet); Plaintiff's Complaint at ¶¶ III.27-.28, Cook v. Ochsner Found. Hosp., 61 F.R.D. 354, 364 (E.D. La. 1972), *reprinted in* THE HOSPITAL, *supra* note 38, at 71 (alleging that two named indigent black residents of New Orleans who suffered from congestive heart failure died after they had been refused care at a public hospital because of lack of facilities, and at three private nonprofit hospitals because of their inability to provide cash deposits ranging from $200 to $250).

43. *See* Randal, *Coronary Artery Bypass Surgery*, HASTINGS CENTER REP., Feb. 1982, at 13, 14-15.

44. *See id.* at 14 (reporting that a consensus panel assembled by the National Heart, Lung, and Blood Institute and the (now defunct) National Center for Health Care Technology in December 1980 and April 1981 could not agree on the value of CABG for many types of patients).

45. *Id.* at 13.

46. *See id.* at 14-15 (reporting the findings of a 1974-1979 Coronary Artery Surgery Study and the National Center for Health Statistics' 1980 Hospital Discharge Survey).

blacks earlier than whites."[47] No studies have been done to explain this apparent underrepresentation of blacks; the obvious explanation is that it is related to the fact that "patients of lower socioeconomic status, which would include many in minority groups, generally have poorer access to the health care system . . . ."[48]

Although Blumstein's argument is hard to follow, there appear to be two reasons for the disappearance of the poor from his analysis of rationing. The first is Blumstein's apparent belief that access to health care by the poor is largely a function of income distribution, and that the subject of income distribution, in turn, is distinct from that of rationing. Thus, in a section entitled "The Vision of Equity in a Market System,"[49] he writes that "[t]he competitive approach . . . does not directly address the question of equity of access, except that it rejects the ideological stances that (1) health is . . . not subject to tradeoffs and (2) equal utilization of medical services is an absolute equitable imperative . . . ."[50] Blumstein's point seems to be that financial barriers to obtaining health care reflect not a problem of the organization or financing of *health care*, but rather a separate problem of *income distribution*. To use a perhaps clearer example, if the poor are unable to purchase automobiles, we are not likely to say that society is engaged in rationing automobiles by price—although in some sense this is true—or that we need a new "automobile policy."

The weakness of this analysis when applied to health care should be readily apparent. For reasons discussed earlier, the health care market has never operated for the poor (or indeed for anyone) in the way that markets operate with respect to other goods and services. The contrast between automobiles and health care is instructive. Advertisements for automobiles fill the daily newspapers and television programs, and car dealers are relatively accessible throughout large metropolitan regions and offer generally similar goods and services. Health care is organized, paid for, and experienced in a very different manner. The extent of patient information, capacity to choose among providers, and ability to pay is far less than in almost any market. Decades of economic and social inequality have resulted in at least two distinct systems of care, one for the poor and minorities, the other for the more affluent, with degrees of variation in between. Access to medical care is thus not simply a function of current income, although that is certainly important. It is also a function of traditions of insurance or governmental coverage, the location and organization of hospitals, physician practices, and patterns of provider and patient expectations as influenced by attitudes about race and social class.[51]

---

47. *Id.* at 15.
48. *Id.*
49. Blumstein, *supra* note 1, at 1351-52.
50. *Id.* at 1352. *See also id.* at 1349. Blumstein does not refer to anyone who asserts the propositions he denies; in fact it is doubtful that any participant in contemporary American health policy debates actually holds such extreme positions.
51. *See* Meyer v. Massachusetts Eye & Ear Infirmary, 330 F. Supp. 1328 (D. Mass. 1971) (doctor prohibited by hospital from informing Medicaid patients about existence of "private" services); R. DUFF & A. HOLLINGSHEAD, *supra* note 19 (documenting differential treatment of patients by physicians according to a patient's social class).

The second way that Blumstein eliminates the distinctive problems of the poor from his analysis is by merging their problems into the issue of "cost containment." When Blumstein refers to "ration[ing] the use of public funds allocated to medical care," the examples he offers include reducing covered services, restricting eligibility, placing an absolute limit or "cap" on government spending, and totally eliminating programs.[52] Although Blumstein does not mention any particular "program," veterans of health care debates have no difficulty identifying most of these examples as coming from the history of the major medical program for the poor, Medicaid.[53] Blumstein characterizes these government policies as "rationing" designed to achieve "cost containment,"[54] thereby implicitly equating them with the goals and techniques of the new rationing.

The equation of current government policies to reduce Medicaid spending with the new rationing is mistaken. All of the decisions Blumstein characterizes as "rationing" and "cost containment" are designed simply to limit government expenditures for the poor without regard for their actual needs—the very hallmark of the traditional rationing practices. His statement that "[g]overnment can reduce the coverage of a program"[55] does not convey what is often at stake. Virginia, for example, reduced its Medicaid coverage by refusing to pay for more than twenty-one days of hospital care per year, regardless of a patient's actual need.[56] New York reduced its coverage by refusing to pay for surgery except in cases of "severe pain" or "immediate" threat of death or disability.[57] Many states have refused to pay for drugs such as L-dopa, which permits victims of Parkinson's disease to engage in basic human activities.[58]

Policies such as these "contain costs" only by limiting the size of the government's health care budget for the poor. They do little or nothing to allocate resources on the basis of need, reduce the unit cost of care, or eliminate wasteful services. Instead of containing costs in any defensible sense of the term, they shift and conceal the impact of the costs. A hospital told that it will no longer be paid by Medicaid for lengthy hospital stays, no matter how medically justified, is likely to pursue two policies. First, it will seek to include its unreimbursed costs in its bills to other patients, if this is permissible under state law and the terms of the hospital's contracts with Blue Cross and other insurance

52. Blumstein, *supra* note 1, at 1356 & n.50.

53. *See, e.g., Health Care Reform*, *supra* note 10, at 292-95.

54. Blumstein, *supra* note 1, at 1356.

55. *Id.* at 1356 n.50.

56. *See* Virginia Hosp. Ass'n v. Kenley, 427 F. Supp. 781 (E.D. Va. 1977).

57. *See* Medical Soc'y v. Toia, 560 F.2d 535 (2d Cir. 1977); Sparer, *Punishing the Victim: A Comment on* Medical Society of New York v. Toia *and the Two Camps of Liberal Health Care Reformers*, 2 HEALTH L. PROJECT LIBR. BULL. 1 (1977).

58. *See* Bass v. Rockefeller, 331 F. Supp. 945 (S.D.N.Y. 1971). *See also* Benton v. Rhodes, [1976-1977 Transfer Binder] MEDICARE & MEDICAID GUIDE (CCH) ¶ 28,025 (S.D. Ohio May 11, 1976) (injunction granted against state reduction in Medicaid coverage that would exclude payment for, inter alia, valium needed to control violent muscle spasms in patients with spinal cord injuries), *rev'd*, 586 F.2d 1 (6th Cir. 1978), *cert. denied sub nom.* Wisebaker v. Rhodes, 440 U.S. 973 (1979).

carriers.[59] Second, the hospital will begin to tighten its admissions and discharge policies by refusing to admit certain types of patients and by discharging patients when their Medicaid coverage expires.[60] Similar problems may exist for all patients whose insurance coverage falls short of their needs, but Medicaid patients are by definition without other financial resources, and the coverage limitations imposed on them by the states are typically far more restrictive than those found in most insurance policies.

III. The Failure of Malpractice Law to Protect the Poor

Blumstein argues that while government has great legal latitude to withhold or discourage beneficial treatment, hospitals and physicians do not. Malpractice law, he maintains, requires hospitals and doctors to deliver needed care unless they can both justify withholding it on the basis of some professional norm (albeit possibly that of a "respectable minority") and obtain the patient's informed consent.[61]

Although Blumstein gives an accurate statement of malpractice principles, he overstates the extent to which malpractice law actually constrains health care rationing, especially for the poor. Malpractice law does not even reach the major devices used to withhold treatment from the poor: those that prevent a doctor-patient relationship from coming into existence. Since malpractice law typically focuses on the fault of a doctor or hospital within an existing patient relationship, it does not cover initial refusals of care, except perhaps in cases of emergency.[62] Other sources of law—notably the federal Hill-Burton statute and regulations,[63] federal and state health planning laws,[64] and state

59. Several states permit hospitals to include the unreimbursed costs of serving the poor in their bills to other patients. *See, e.g.,* N.J. STAT. ANN. § 26:2H-18(d) (West Supp. 1981-1982).

60. *See* Sparer, *supra* note 10, at 10-11.

61. Blumstein, *supra* note 1, at 1395-99.

62. *See, e.g.,* J. KING, THE LAW OF MEDICAL MALPRACTICE 22 (1977). Most states do not impose an affirmative duty on doctors to render emergency treatment, but many do recognize such a duty on the part of hospitals that maintain emergency care facilities. *Compare* Hurley v. Eddingfield, 156 Ind. 416, 59 N.E. 1058 (1901) (physician may refuse request for emergency care, even though fee tendered), *with* Wilmington General Hosp. v. Manlove, 54 Del. 15, 174 A.2d 135 (1961) (hospital has duty to render emergency care to patients who have relied on its custom of providing emergency services), *and* Guerrero v. Copper Queen Hosp., 112 Ariz. 104, 106, 537 P.2d 1329, 1331 (1975) (state licensing regulations requiring hospitals to maintain emergency room services create duty on hospital "not [to] deny emergency care to any patient without cause"). *See also* ILL. ANN. STAT. ch. 111½, § 86 (Smith-Hurd 1977) (licensed hospitals required to furnish emergency service to any applicant whose condition is liable to cause death, severe injury, or serious illness); VT. STAT. ANN. tit. 12, § 519 (1973) (unusual state law imposing an affirmative duty on all persons, including physicians, to give "reasonable assistance" to another "exposed to grave physical harm" if the assistance can be rendered "without danger or peril to himself" and without interference "with important duties owed to others"); AMERICAN MEDICAL ASS'N CODE OF ETHICS § VI (physician is free to choose whom to serve "except in emergencies"), *reprinted in* N.Y. Times, July 23, 1980, at 12, col. 2).

63. *See* Hospital Survey and Construction (Hill-Burton) Act, 42 U.S.C. § 291c(e) (1976); 42 C.F.R. §§ 124.1-.607 (1981) (promulgating new final regulations on provision of service to persons unable to pay, and on required "community service" by federally assisted health facilities).

64. *See* National Health Planning and Resources Development Act of 1974, Pub. L. No. 93-641, 88 Stat. 2225 (codified as amended in scattered sections of 42 U.S.C.). The relationship of this statute to the health needs of the poor is discussed extensively in Schneider & Wing, *The National Health Planning and Resources Development Act of 1974: Implications for the Poor*, 9 CLEARINGHOUSE REV. 683 (1976). *See also Health Care Reform*, *supra* note 10, at 304-30. For analogous state legislation, *see, e.g.,* New Jersey Health Facilities Planning Act of 1971, N.J. STAT. ANN. § 26:2H-1 (West Supp. 1981-1982).

licensing standards[65]—incorporate in various ways the value of access for needed care. But in the face of industry opposition, the agencies and courts responsible for interpreting and enforcing these laws have almost never done so in ways that benefited low-income patients.[66]

Malpractice law has also not proven effective at regulating the rationing of care for the poor within existing physician-patient relationships. Private, nonprofit hospitals have been known to discharge or "dump"[67] patients when it becomes clear that the costs of their care will exceed their ability to pay. Both the professional and legal norm in these cases is clear: a patient should not be discharged or "abandoned" without reasonable notice to the patient and sufficient time to obtain an alternative source of care.[68] Despite this well-known norm, there is reason to believe that it is frequently violated. The records of Cook County Hospital in Chicago, for example, contain many examples of patients being sent from other hospitals—sometimes hospitals in other states—without adequate records, with no consultation with the receiving hospital, and in conditions that should have precluded any movement.[69] In Philadelphia, a patient suffering from advanced emphysema and needing oxygen and other assistance was evicted from a private hospital because he had exhausted sixty days of hospital coverage under Medicaid.[70] He died before he could benefit from a consent decree.[71]

Another form of rationing for the poor within the doctor-patient relationship arises from the substandard conditions in many public hospitals. Legislative and judicial proceedings in numerous cities, including Washington, D.C., Philadelphia, New York, Newark, and San Francisco, have repeatedly uncovered such dangerously substandard

65. *See, e.g.,* CAL. HEALTH & SAFETY CODE § 1317 (West 1979); ILL. ANN. STAT. ch. 111½, § 86 (Smith-Hurd 1977).

66. *See, e.g., Health Care Reform*, *supra* note 10 (analyzing litigation seeking agency enforcement of Hill-Burton, Medicaid, and health planning statutes and regulations); Rose, *Federal Regulation of Services to the Poor Under the Hill-Burton Act: Realities and Pitfalls*, 70 NW. U.L. REV. 168 (1975) (documenting long history of federal and state failure to enforce the redistributive requirements of the Hill-Burton Act); Sparer, *Health Planning For—or Against—Innovative and Improved Maternity Care*, 5 HEALTH L. PROJECT LIBR. BULL. 291 (1980) (documenting failure of local Health Systems Agency to enforce both federal and local guidelines on access to innovative childbirth services); Sparer, *The Legal Right to Health Care: Public Policy and Equal Access*, HASTINGS CENTER REP., Oct. 1976, at 39.

67. The term "patient dumping" has been used to refer to two somewhat different, although related, phenomena. The first is the meaning stated in the text: the *termination* of an existing hospital-patient relationship because of the patient's inability to pay. *See* Sparer, *supra* note 10, at 10. The second is the practice of *blocking access* for the poor by refusing to provide outpatient services and other points of entry. *See* Roemer & Mera, *"Patient-Dumping" and Other Voluntary Agency Contributions to Public Agency Problems*, 11 MED. CARE 30 (1973).

68. *See* J. KING, *supra* note 62, at 22-27.

69. About 500 patients per year are transferred to Cook County Hospital because of their inability to pay the costs of treatment at other hospitals. Many of these transfers are medically inappropriate or not properly arranged with the receiving hospital. *See* Sparer, *supra* note 10, at 11.

70. *See* Turner v. Wohlgemuth, No. 74-1680 (E.D. Pa. 1976) (consent decree), *discussed in* Sparer, *supra* note 10, at 10 n.28.

71. According to Professor Sparer, this case led the Pennsylvania Legislature to repeal the 60-day hospitalization limit. *Id.*; *see* Act of July 15, 1976, Pub. L. No. 202, § 6, 1976 Pa. Laws 993, 1006, *amending* Act of July 31, 1968, Pub. L. No. 273, § 5, 1968 Pa. Laws 904, 907.

conditions as emergency rooms staffed by interns,[72] inadequate nursing in intensive care,[73] lack of vital equipment and supplies,[74] and inadequate medical records and laboratory systems,[75] all of which can increase disease and death among patients.[76] Although in one case a federal district court issued an injunction against such conditions on the grounds that they constitute a common-law nuisance and jeopardize the patients and the community that relies on the hospital,[77] such a decision is very rare,[78] and the conditions persist in many cities. Despite the obvious violation of professional and legal standards of care, cities continue to underfund their facilities, apparently with little fear of malpractice liability.

Have malpractice standards really provided inadequate protection to the poor? There is no systematic study of this question, but some available data suggest the absence of effective malpractice remedies. A study that analyzed 1970 data for the Secretary of Health, Education, and Welfare (HEW) Malpractice Commission found that in a sample of malpractice cases where the insurance coverage of the parties was known, only 1.2% of the malpractice claimants were covered by Medicaid.[79] In light of the probable incidence of malpractice among the 14.5 million Medicaid recipients who, in 1970, received over $1.8 billion in inpatient hospital services,[80] this is a remarkably low percentage. The private hospitals' discharge practices and public hospitals' substandard conditions have been widely documented, and one would expect them to lead to compensable injuries.[81] Moreover, when Medicaid patients do gain access to medical and hospital care, they are twice

72. *See* Greater Washington, D.C. Area Council of Senior Citizens v. District of Columbia, Civ. Act. No. 275-71 (D.D.C. Sept. 27, 1971), *reprinted in* THE HOSPITAL, *supra* note 38, at 434 (District of Columbia General Hospital found to have violated its legal obligation to provide adequate emergency care in accordance with its own rules and accepted medical practice by staffing its emergency room only with interns).

73. *See* Greater Washington, D.C. Area Council of Senior Citizens v. District of Columbia, 406 F. Supp. 768, 773 (D.D.C. 1975); Boone v. Tate, No. 713 (C.P. Philadelphia County, Apr. 30, 1971), *reprinted in* THE HOSPITAL, *supra* note 38, at 440, 445.

74. *See* Greater Washington, D.C. Area Council of Senior Citizens v. District of Columbia, 406 F. Supp. 768, 773 (D.D.C. 1975) (discussing shortages of drugs and breakdowns of X-ray machines).

75. *See id.* at 770-71, 774.

76. *See id.* at 772 (lack of emergency room staff "contributed to the death of several patients").

77. *Id.* at 775-76.

78. The rarity of this decision is suggested by the fact that not a single federal or state court has cited it in a published opinion since its issuance in 1975. For a contrary judicial ruling finding no municipal duty to remedy dangerously substandard hospital conditions, *see* Boone v. Tate, No. 713 (C.P. Philadelphia County, Apr. 30, 1971), *reprinted in* THE HOSPITAL, *supra* note 38, at 440. Extensive judicial regulation of substandard conditions in public mental hospitals has occurred because of the need to justify confinement. *See, e.g.,* Wyatt v. Stickney, 344 F. Supp. 373 (M.D. Ala. 1972), *aff'd*, 503 F.2d 1305 (5th Cir. 1974).

79. *See* Rudov, Myers & Mirabella, *Medical Malpractice Insurance Claims Files Closed in 1970*, in SECRETARY OF HEALTH, EDUCATION & WELFARE'S MALPRACTICE COMMISSION, REPORT ON MEDICAL MALPRACTICE app. at 1, 11 (1973) [hereinafter cited as REPORT ON MEDICAL MALPRACTICE].

80. HEALTH CARE FINANCING ADMINISTRATION, THE MEDICAID BUREAU, U.S. DEP'T. OF HEALTH, EDUCATION & WELFARE DATA ON THE MEDICAID PROGRAM FISCAL YEARS 1966-77, at 29, 32 (1977).

81. *See, e.g.,* Greater Washington, D.C. Area Council of Senior Citizens v. District of Columbia, 406 F. Supp. 768, 768 (D.D.C. 1975) (documenting numerous examples of substandard conditions and care); Sparer, *supra* note 10, at 10-11 (describing harmful impact of exclusionary private hospital policies).

as likely to be subjected to surgery than other patients, thereby increasing the risk of malpractice.[82] Doctors tend to treat patients from lower socioeconomic classes with less respect, less concern for their autonomy and consent, and less emphasis on effective communication, than extended to higher status patients.[83] Such practices increase the risk of poor diagnosis and treatment, and themselves constitute a form of malpractice.[84]

Assuming that the HEW Malpractice Commission study accurately reflects litigation patterns, why do so few Medicaid recipients file malpractice claims? For the reasons suggested above, the answer is unlikely to be that they are generally receiving better quality care than the population at large. Again, no systematic study of this question has been done, but some reasonable speculations are possible. Malpractice litigation is expensive and largely financed by contingent fees. Lawyers in private practice thus will not undertake a case unless it has a substantial dollar "worth."[85] But the standard criteria of economic loss from medical injury—lost wages, lost earning capacity, and medical expenses—are not easily applied to patients whose income comes from welfare payments, and whose medical care is paid for—however inadequately—by the state.[86] The most accessible attorneys for the poor are thus likely to be legal services lawyers who are not dependent on fees but are prohibited by statute from accepting the types of cases—such as medical malpractice—that generate fees for the private bar.[87]

In addition, low-income patients may have low expectations of how they should be treated by institutions and professionals and little tradition of contact with the malpractice bar. Patients on welfare have little incentive to seek recovery, "because welfare departments in all states require that, as a condition of receiving public aid, the poor person sign over to the welfare department any money that he or she receives from a personal-injury action."[88] Causation may also be difficult to prove. In many instances, substandard hospital conditions and patient discharge policies do not directly cause death or major injury, but rather increase the risk of their occurrence. Malpractice law is not well designed to compensate for, or to deter, such increased risk for any patients, including the poor. In other cases, serious injury may occur, but it may be difficult to prove that an already-ill patient was made worse by a specific act or hospital condition. Finally, malpractice standards and litigation typically focus on a particular physician's behavior and whether the physician deviated from a professional norm,

82. *See* S. Law & S. Polan, Pain & Profit: The Politics of Malpractice 19 (1978).

83. *See* R. Duff & A. Hollingshead, *supra* note 19.

84. On the relationship of poor communication to incorrect diagnosis and treatment, *see id.* at 151-78; Note, *Restructuring Informed Consent*, 79 Yale L.J. 1533 (1970). For arguments that physicians' failure to respect patients' autonomy not only contributes to worse results in traditional medical terms, but also undermines human capacities essential to healing more broadly conceived, see Lander, *Doctor-Patient Models: Reformist Dilemmas*, 4 Health L. Project Libr. Bull. 2, 5-6 (1979).

85. *See* S. Law & S. Polan, *supra* note 82, at 84 (studies indicate that most New York attorneys will not institute a malpractice action unless the recovery is likely to exceed $25,000-40,000).

86. *Id.* at 20.

87. *See* 42 U.S.C. § 2996f(b)(1) (1976 & Supp. IV 1980).

88. S. Law & S. Polan, *supra* note 82, at 20.

while hospital discharge practices and substandard conditions are usually the responsibility of many people operating in an institutional framework. Responsibility, therefore, is harder to prove, and often knowledgeable witnesses will not testify against institutional practices. In light of all of these factors, it is not surprising that malpractice law has not significantly constrained hospitals and doctors from withholding care from the poor.

IV. Rationing, Excessive Health Services, and Malpractice Law

While the poor often lack access to needed care, they, like middle class patients, are also apt to be victims of overuse of diagnostic tests, drugs, hospitalization, and surgery.[89] Financial and professional incentives toward overuse of costly health care resources have in turn generated the current interest in health care rationing. Many physicians believe, however, that they will be penalized by malpractice law if they provide less than the excessive norm.[90] Professor Blumstein apparently agrees that this belief accurately reflects the law, for he argues that

> a provider who seeks to ration medical care resources—for example, by ordering fewer or less costly tests, by providing a smaller margin of safety in terms of facilities or equipment availability, or by allowing less highly credentialed personnel to perform certain procedures—runs the risk of increased exposure to malpractice liability . . . .[91]

This argument is flawed by its abstraction. The extent to which malpractice law does or might constrain rationing depends upon two factors: the specific rationing at issue, and the sources and content of malpractice law itself. The examples offered by Blumstein concerning tests, facilities, and personnel might or might not result in an increased risk of liability, depending on critical but unstated circumstances. For example, some rationing schemes for withholding tests, equipment, and qualified personnel have departed so far from current professional

---

89. *See, e.g.*, S. LAW & S. POLAN, *supra* note 82, at 18-19; Sparer, *supra* note 10, at 11-14; *Health Care Reform*, *supra* note 10, at 244 nn.5-7. A recent article in the *New York Times* provided the following "vivid example":

> An elderly patient was admitted to the hospital late one night this winter, suffering from bronchitis and running a fever. In the judgment of . . . an experienced internist who was not involved in treating the woman and [who] requested that his name not be used, her illness could have been safely and effectively treated at home with antibiotics.
>
> Undoubtedly, the doctor said, initial X-rays and an initial electrocardiogram would have been justified in view of the woman's age. After that she could have been started on the antibiotics, kept in the hospital for a day or two and sent home.
>
> Instead, she was kept in the hospital for nine days. She was given $575 worth of intravenous infusions, $657 worth of electrocardiograms, $1,454 worth of X-rays and $944 worth of laboratory services. The total hospital bill came to $9,065.10, nearly all of it paid by Medicare. No fraud was involved, according to the physician. Rather, the doctors were simply doing what they were trained to do as part of a profession that has become enthralled with technology.

Stevens, *supra* note 22, at 50, cols. 3-4.

90. The Commission specifically inquired whether this was an accurate reflection of the law. *See* Baily letter, *supra* note 21, at 2. For discussions of physician attitudes, see S. LAW & S. POLAN, *supra* note 82, at 114-19; Bernzweig, *Defensive Medicine*, in REPORT ON MEDICAL MALPRACTICE, *supra* note 79, app. at 38, 40; Carlson, *A Conceptualization of A No-Fault Compensation System for Medical Injuries*, 7 LAW & SOC. REV. 329, 334-35 (1973) (reporting physician attitudes with little documentation or critical evaluation); *The Medical Malpractice Threat: A Study of Defensive Medicine*, 1971 DUKE L.J. 939, 941-43 [hereinafter cited as *1971 Duke L.J. Malpractice Study*].

91. Blumstein, *supra* note 1, at 1395.

norms that they clearly and properly constitute malpractice. The major examples have occurred in public hospitals for the poor and in some of the prepaid health plans that provided low-quality service to Medicaid recipients in California.[92] Unfortunately, for reasons already discussed, malpractice law has not functioned as an effective deterrent to this type of rationing.

On the other hand, many contemporary efforts to reduce excessive testing, hospitalization, and surgery fall well within existing professional norms and seem to pose no risk of increased malpractice liability. Blumstein's reference to "providing a smaller margin of safety in terms of facilities or equipment availability"[93] obscures the issue; one cannot know whether providing less costly facilities and equipment expands or contracts the "margin of safety" without considering specific examples. A properly designed childbirth program, for example, might encourage most women to give birth out of the hospital, either in birthing centers or at home, with appropriate hospital backup. The major providers of such a service would probably be certified-nurse midwives. In this case, providing care with less expensive facilities and less "highly credentialed" personnel would not be likely to expand malpractice liability. On the contrary, there is reason to believe that such a program would provide better quality care with less risk of over-intervention and related malpractice liability.[94] The major reason why few of these childbirth programs exist is not an increased risk of malpractice liability, but rather pervasive, often subtly manifested opposition of physicians and hospitals who benefit from current childbirth practices.[95]

Malpractice law could—at least in theory—accommodate a substantial amount of health care rationing because the law is largely within the control of the medical profession itself. This possibility is apparently widely rejected by physicians, who often perceive themselves as conscientious professionals assailed by unreasonable legal standards. For example, in a survey undertaken by the American Medical Association in 1970, seventy percent of the responding physicians stated that they practiced defensive medicine, meaning that at least some of their decisions were based not on clinical judgments but on fears of legal liability.[96] For several reasons, however, it seems likely that this view is mistaken, and that the often-blamed phenomenon of "defensive medicine" is largely an excuse for practices rooted in other motivations.

First, it has been extremely difficult for both the medical and legal professions to define defensive medicine or measure its extent and sig-

92. For a discussion of public hospitals, see sources cited *supra* notes 72-77. On the California experience with prepaid health plans for the poor, see Schneider & Stern, *Health Maintenance Organizations and the Poor: Problems and Prospects*, 70 Nw. U.L. Rev. 90, 126-38 (1975).

93. Blumstein, *supra* note 1, at 151.

94. *See* Levy, Wilkinson & Marine, *Reducing Neonatal Mortality Rates with Midwives*, 109 Am. J. Obstetrics & Gynecology 50 (1971); Mehl, *Home Delivery Research Today—A Review*, 1 Women & Health — (1976); Stimeling, *Will Common Delivery Techniques Soon Become Malpractice?*, J. Legal Med., May 1975, at 20; *cf. Market Failures*, *supra* note 16, at 1096-99.

95. *See* Sparer, *supra* note 66; *Inadequate Childbirth Care: A Different Perspective on Law and Lawyers*, 3 Health L. Project Libr. Bull. 5, 8-9 (1978).

96. *See* American Medical Association, 1970 Professional Liability Survey, *discussed in* S. Law & S. Polan, *supra* note 82, at 115; Bernzweig, *supra* note 90, at 39-40.

nificance.[97] According to dominant medical opinion, many tests and procedures provide some additional knowledge or possible benefit, and therefore are probably not motivated solely by fear of legal liability.[98] One of the few careful studies on the subject identified over three dozen tests and procedures widely suspected of being performed because of fear of liability, and then surveyed physicians in North Carolina and California to ascertain their motivations for undertaking these procedures.[99] The study concluded that while fear of litigation does cause some overuse of tests and other procedures, "the practice [of defensive medicine] is not extensive and [is] probably not a contributing factor to the rising costs of medical care."[100] A background paper for the HEW Malpractice Commission came to a similar conclusion, stressing as causal factors the "lack of cost constraints" and the laboratory emphasis of modern medical education.[101]

Second, with rare exceptions, the legal standard of care demanded of physicians is that of the medical profession itself.[102] To win a malpractice suit, an injured patient must normally show by expert medical testimony that the treating physician violated the profession's own standards of what a reasonable doctor would have done in similar circumstances.[103] Moreover, malpractice law does not assume that the medical profession is monolithic; as Blumstein points out, a physician will not be penalized for following the practices of a "respectable minority."[104] Given the general deference of malpractice law to the medical profession and the additional possibility of obtaining the patient's informed consent to practices that differ from the excessive norm, a rather wide scope would appear to exist for rationing programs that have some basis in professional knowledge and honestly inform patients about relevant risks and benefits.

Although the principles of malpractice law, properly understood, are receptive to justified rationing, physicians' beliefs about malpractice law are not. Physicians' attitudes or fears about malpractice are often inconsistent with, or at least not verified by, the actual workings of the malpractice system.[105] Nevertheless, these views are deeply rooted in the self-interest of the medical profession, and they will probably remain influential. As Professor Peter Schuck correctly points out, the majority of the medical profession is unlikely to develop norms that require or actively encourage rationing.[106] Significant changes in diagnosis and treatment practices will, in all likelihood, be neither pro-

97. *See* Bernzweig, *supra* note 90, at 38-39; *1971 Duke L.J. Malpractice Study*, *supra* note 90, at 943.

98. *See* Bernzweig, *supra* note 90, at 39.

99. *See 1971 Duke L.J. Malpractice Study*, *supra* note 90, at 953-54, 966-67.

100. *Id.* at 964.

101. Bernzweig, *supra* note 90, at 40.

102. *See* S. LAW & S. POLAN, *supra* note 82, at 7; Capron, *Tort Liability in Genetic Counseling*, 79 COLUM. L. REV. 618, 667 (1979).

103. *See* Capron, *supra* note 102; Blumstein, *supra* note 1, at 151.

104. Blumstein, *supra* note 1, at 152.

105. *See* text accompanying notes 90 & 96-104 *supra*. *See also* S. LAW & S. POLAN, *supra* note 82, at 114-119.

106. *See* Schuck, *Malpractice Liability and the Rationing of Care*, 59 TEXAS L. REV. 1421 (1981).

moted nor inhibited by malpractice law itself, but will rather depend on major social decisions about how consumers are to pay for health care and how providers are to be organized and reimbursed.

V. Conclusion

There is a widespread consensus among knowledgeable observers that the current American health care system is both wasteful and inequitable.[107] The waste is inherent in a system where quality is defined in terms of high technology, treatment and investment decisions are largely unregulated, and provider reimbursement and patient insurance policies weaken concern with costs. The inequity is inherent in a system that fails to assure basic health care coverage for many low-income and unemployed persons or produce an adequate supply of, and access to, needed types of providers.

In this context, rationing of health care services is urgently needed. Not surprisingly, however, it has been pursued in what is perhaps the most inequitable manner possible: by withholding needed and minimally adequate care from many low-income patients, while subjecting middle-income (and some poor) patients to excessive services that are often harmful as well as costly. The various sources of American law—the Constitution, statutes, regulations, and judge-made malpractice law—have either not addressed these issues, or have largely acquiesced in the existence of waste and inequity. Given this background, rationing efforts should address questions of social class and professional power. As with most important issues, American law provides some concepts and values to assist in grappling with these questions, as well as a long tradition of denying their importance and legitimacy.

107. The literature supporting this point is vast, and spans the political spectrum. *See, e.g.*, A. Enthoven, Health Plan (1980); Arras, *Health Care Vouchers for the Poor*, 11 Hastings Center Rep., Aug. 1981, at 29; Sparer, *supra* note 10. A recent public opinion poll by the *New York Times* suggests that the general public partially agrees with this consensus. On the one hand, the poll found Americans to be "remarkably pleased with their basic medical care, despite certain pockets of dissatisfaction, particularly among minorities and the poor and despite numerous unhappy experiences with some doctors." Reinhold, *Majority in Survey on Health Care are Open to Changes to Cut Costs*, N.Y. Times, Mar. 29, 1982, at A1, D11:1. At the same time, the survey found extensive concern about the price of medical care, and a majority or "sizable numbers" willing to alter substantially their relationship with providers in order to reduce costs. *Id.* Other studies suggest that surveys reporting patient satisfaction should be viewed with care. *See* Sparer, *supra* note 10, at 14 nn.53-54.

Malpractice Liability and the Rationing of Care

Peter H. Schuck*

The current relationship of malpractice law to decisions to ration health care may readily be summarized: unless it can be shown that at least a "respectable minority" of practitioners in the relevant, increasingly national geographical area would make a similar rationing decision under similar circumstances, a decision to ration will expose the practitioner to the risk of malpractice liability. I use the phrase "unless it can be shown" instead of "unless the plaintiff can show" because the legal principle of res ipsa loquitur will, under certain circumstances, create an inference that the physician was negligent even in the absence of direct proof of negligence.[1] Once created, this inference shifts the burden of explanation from the patient to the physician. As a practical matter, if not as a legal one, it may shift the burden of proof as well.

I would add one other point to this summary. Professor Blumstein correctly observes that the professional standard is almost always the standard of care to which the physician will be held. In *Helling v. Carey*,[2] however, the one well-known case in which a court refused to be bound by the professional norm, the norm that the court rejected actually *favored*, or at least reflected, a rationing decision, while the norm that the court imposed in its place *rejected* rationing. Thus, not merely do professional norms tend to eschew rationing, but when a

* Professor, Yale Law School. B.A. 1962, Cornell University; J.D. 1965, Harvard University; LL.M. 1966, New York University; M.A. 1969, Harvard University.

This paper was originally presented to the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research, which had asked the author to enlarge upon Professor Blumstein's analysis of the effect of malpractice law on health care rationing. The author assumes for purposes of discussion that public policy might wish to encourage such rationing, and that the problem is whether and how this can be accomplished in view of the risk of malpractice liability that health care providers can encounter.

1. W. Prosser, The Law of Torts § 39 (4th ed. 1971).
2. 83 Wash. 2d 514, 519 P.2d 981 (1974).

Published in 59 Texas L. Rev. 1421 (1981). Copyright © 1981 by the Texas Law Review. Reprinted by permission.

court musters the hubris to criticize a professional norm, it is *plus royaliste que le roi*—even more anti-rationing than the doctors. In short, if rationing is desirable, it is probably fatuous to expect courts in malpractice cases to legitimate it without the benefit of a statute.[3]

It is most unlikely that malpractice standards based upon professional norms favoring rationing will emerge unaided from the professional birth canal; induced labor will be necessary for several reasons. First, for the profession itself to recognize an ethical duty to ration care would fundamentally challenge the very meaning of its professionalism—the physician's unequivocal, undivided loyalty to the patient's best interests. The fact that the physician often defines the patient's interests, and that the physician's professional and financial interests may sometimes conflict with those of the patient, only underscores the subversion of professional values. Each of these facts tends to strengthen, rather than weaken, the profession's hostility to rationing.

Second, quite apart from the dictates of professionalism, it would be irrational for an individual physician to ration care, even if he or she believed that rationing constituted sound social policy and would, if generally followed, ultimately benefit patients. The reason is that an individual physician's decision to ration would have no impact upon the behavior of other physicians or health care providers. The physician and patient would bear the cost of the rationing decision without receiving any of the benefits (except for escaping a copayment charge) that a generalized practice of rationing might bestow.

Third, for a norm favoring rationing to have any practical effect, it could not take the form of a general exhortation to ration; rather, it would have to be specific enough to provide guidance to individual providers making particular treatment decisions. To be that specific, however, the norm could not merely affect "statistical" lives but would have to target specific groups of patients, if not "identifiable" lives.[4] Professor Blumstein correctly argues that society resists making explicit allocative decisions about the latter.[5] But the medical profession would find such decisions even more morally repugnant than would society in general.[6]

The rationing of health care in our society can be legitimated in only a few ways. Consumers themselves ration through their consumption decisions. Health care providers ration by selecting patients; for example, by refusing to participate in Medicaid they force patients to seek care elsewhere or, perhaps, to forego it altogether. Society rations by limiting public health care expenditures, imposing licensing restric-

3. In fact, the result in *Helling* was quickly reversed by the Washington Legislature, which restored professional standards as the only criterion of reasonable care in malpractice cases. Act of May 12, 1975, ch. 36, 1975 Wash. Laws. 252 (codified at WASH. REV. CODE § 4.24.290 (Supp. 1981)). A small step for rationing; a giant leap for professional autonomy.

4. Sometimes those restrictions, although implemented at the macro level, can affect lives that are reasonably "identifiable." Examples would include a prohibition on Medicaid funding of abortions or a limitation upon Medicare reimbursement for end-stage renal disease victims.

5. Blumstein, *Rationing Medical Resources: A Constitutional, Legal, and Policy Analysis*, 59 TEXAS L. REV. 1345, 1353-54 (1981); *see, e.g.*, G. CALABRESI & P. BOBBITT, TRAGIC CHOICES (1978).

6. *See* Scheff, *Decision Rules, Types of Error, and their Consequences in Medical Diagnosis*, 8 BEHAV. SCI. 97 (1963).

tions on health care personnel, and giving legal effect to compliance with Professional Standard Review Organization (PSRO) norms, which confer immunity from malpractice liability.[7]

Significantly, each of these three types of rationing decisions ultimately depends upon the consent of patients, either as a class or as individuals, for its legitimacy. Social rationing decisions are legitimate only to the extent that they emerge from democratic political processes in which the consent of the public—including the large group of people who may expect to become patients—is obtained. Health care provider decisions that ration by refusing to enter into a provider-patient relationship with a particular individual may be validated in terms of the values of professional autonomy and freedom of choice. But once such a relationship has been established and the question is whether or how to provide treatment, only the patient's informed consent can legitimate further rationing. Legitimation of the personal rationing decisions of the consumer depends upon the nature of the rationing. If the question is whether to seek care at all, the legitimacy of the decision turns upon our willingness to accept the distribution of wealth, opportunity, and information that constrains individual choice. But if the question concerns what care is to be consumed, then the legitimacy of the decision depends not only upon the underlying distributional questions of purchasing power and access, but also upon the patient's informed consent.

I have stressed the importance of consent not only because it constitutes a fundamental ethical value in health care, but also because it suggests several ways in which the professionally dominated standards of malpractice law might be accommodated to the economizing values of rationing decisions. The first of these approaches would encourage rationing by altering the *substantive norms* of conduct applied in malpractice cases. The second would encourage rationing by altering the *identity* of the provider rather than the nature of substantive behavior. The third would encourage rationing by altering the *structure of financial incentives* within which treatment decisions are made.

1. Altering Substantive Norms.—Professionally derived or court-imposed standards of care, as we have seen, are not likely to impose a duty upon the individual provider to ration care. But such norms could be designed to reflect rationing decisions legitimated in some other way. The standards developed by PSROs are examples of this.[8] From the point of view of rationing, the strength of PSRO norms is that they are legitimated by legislation and developed at a level remote from individual treatment and "identifiable lives" and through a process strongly biased against "under-care." Their limitations, of course, are the obverse side of the coin. Because the norms are developed exclusively by local providers who face no financial constraints or incentives to economize, PSROs are likely, as Professor Blumstein has argued

7. *See* Blumstein, *supra* note 5, at 1345, 1356-57, 1388-90, 1397.
8. *See id.* at 1397-98.

elsewhere, to disdain any "quality"-cost trade-off.[9] Their norms, far from encouraging rationing, will simply validate the profession's most expensive forms of practice.

It may be possible, however, to strike a better balance between professional autonomy, fidelity to the patient's interest in optimal care, and the hypothesized public interest in health care rationing. Norms could be developed by groups that include at least some members who are, for institutional reasons, more sensitive to cost considerations than fee-for-service physicians are likely to be. In contrast to PSROs, the norms could be developed not only locally but at the state and national levels where most health care is financed, thereby "internalizing" the costs as well as the benefits generated by the norms. One model for this approach is a National Institute of Health effort to develop national norms of care based upon a negotiated consensus among health care providers, regulators, financers, and consumers.[10] Courts would certainly accept such norms as strong evidence of the applicable standards of care in malpractice cases. If these norms truly reflected so broad a consensus, they might even come to be regarded as conclusive.

2. Changing the Nature of the Provider.—Even where rationing by means of substantive norms of care is unacceptable for one reason or another, rationing provides that the "same care" from a less expensive provider may be ethically acceptable and consistent with professional standards of care. There is solid evidence, for example, that properly supervised auxiliary personnel can render routine dental and medical care with little or no deleterious effect and with equivalent patient satisfaction.[11] Yet self-serving professional practice laws, reimbursement arrangements that favor highly credentialed personnel, prohibitions against institutional licensure, and other impediments now inhibit even this type of rationing. If such restrictions were narrowed or abandoned, innovative and efficient deployments of personnel might well be encouraged without the need to ration particular services and without an increase in malpractice liability.

3. Expanding Financial Incentives.—The current system of third-party, first-dollar insurance not only discourages patient self-rationing; it also renders such behavior economically irrational. The fee-for-service system has much the same effect upon treatment decisions by physicians. Malpractice law mirrors and further entrenches the noneconomizing patterns of care generated by these systems.

To the extent that current policy addresses this problem, it does so by seeking to encourage the proliferation of Health Maintenance Organizations (HMOs) and other organized settings in which health care providers have some incentive and opportunity to ration.[12] There are

9. *See* Blumstein, *The Role of PSROs in Hospital Cost Containment*, in HOSPITAL COST CONTAINMENT 461 (M. Zubkoff, I. Raskin & R. Hanft eds. 1978).

10. *See* Perry & Kalberer, *The NIH Consensus-Development Program and the Assessment of Health-Care Technologies*, 303 NEW ENG. J. MED. 169 (1980).

11. *See* CONGRESSIONAL BUDGET OFFICE, PHYSICIAN EXTENDERS: THEIR CURRENT AND FUTURE ROLE IN MEDICAL CARE DELIVERY (1979).

many theoretical and practical difficulties with the HMO strategy, and it has so far met with only limited success.[13] One problem is that the provider has an incentive to ration, at least to the point that malpractice liability becomes a threat, but the subscriber does not. Indeed, once enrolled, the subscriber faces no constraints on consumption other than whatever queuing costs and physician controls on access may exist. Another difficulty is the evident preference of most Americans for fee-for-service care, at least under current institutional arrangements.

An alternative to the HMO strategy, now under active consideration, is the so-called "competition" strategy. This approach to rationing relies heavily upon the prospect of HMO growth, but it would also require several other reforms, such as eliminating the substantial tax benefits now enjoyed by purchasers of health insurance. These reforms face awesome political difficulties. A competition strategy, if successfully implemented, could encourage groups of competing health care providers to offer different benefit packages reflecting, among other things, different rationing-limited liability combinations. But the adoption and implementation of this strategy seems so problematical that I shall not dwell on it.[14]

There is a less ambitious but perhaps ultimately more promising approach to encouraging rationing in the face of malpractice remedies, at least in the short run. This strategy emphasizes risk-sharing by primary care physicians and would build incrementally upon existing health care institutions, especially upon the physician-patient relationship characterized by fee-for-service care. It requires a system of third-party reimbursement that creates incentives for both physicians and patients to economize yet fully respects and is legitimated by the consent principle. Such systems have been operating since 1975 in three Western States under the administration of SAFECO Insurance Company.[15] The basic concept is quite straightforward: a patient chooses one primary care physician who then acts as a gatekeeper to secondary and tertiary care for that patient. The physician takes responsibility for the coordination and financial management of the patient's care, making all referrals to specialists and hospitals as well as authorizing all payments for care provided to the patient from the physician's own account with the insurer. At the end of the accounting period, the physician shares with the patient in any deficit or surplus remaining in the account.

If we wish to encourage rationing without thereby increasing the risk of malpractice liability, this system has several advantages. Unlike ordinary fee-for-service practice reimbursed by conventional third-party insurance, both patient and health care provider would have a

12. *See, e.g.,* Health Maintenance Organizations Act, 42 U.S.C. §§ 300e-2 to 300e-4 (1976 & Supp. III 1979) (authorization for grants, contracts, and loans to aid in the initial stages of planning, developing, and operating HMOs).

13. *See, e.g.,* Brown, *Competition and Health Cost Containment: Cautions and Conjectures*, 59 MILBANK MEMORIAL FUND Q.: HEALTH & SOC'Y 145 (1981).

14. For a discussion of this issue, see a recent symposium, *Market Oriented Approaches to Achieving Health Policy Goals*, 34 VAND. L. REV. 849 (1981).

15. *See, e.g.,* Moore, *Cost Containment Through Risk-Sharing by Primary-Care Physicians*, 300 NEW ENG. J. MED. 1359 (1979).

palpable incentive to make economizing trade-offs. Unlike most HMO policies, these rationing decisions would be made, if at all, only within the context of a personal, well-defined relationship between the patient and his or her own physician, not in an impersonal, bureaucratic milieu. Thus, the rationing decision is likely to be more focused, visible, and identifiable; the consent of the patient to that decision may be better informed, more genuine, and, therefore, constitute a more viable defense in a malpractice action.

While the risk of "under-care" remains in this system, this risk is not obviously greater than that created by HMO practice, and it may in fact be less. The participating physician's incentive to undertreat or underrefer may be somewhat greater than that of the salaried physician in an HMO; on the other hand, the *in terrorem* effect of the malpractice remedy may on balance be strengthened in a system in which the physician's potential conflict of interest is no longer concealed but is obvious to all concerned, and in which the trade-off between the costs and benefits of a rationing decision is more clearly delineated to the patient.

Professor Blumstein has raised and answered some important questions about how our society has organized itself to make difficult but essential decisions about core human values. He has not blinked at the harsh necessity of choosing, and neither can we. I have argued, however, that the common law of malpractice is a singularly unpromising source of guidance about how we should go about making such choices. I have tried to suggest several ways in which that task might better be structured and performed.

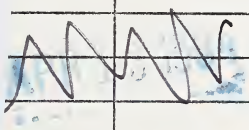
*U.S. GOVERNMENT PRINTING OFFICE : 1983 0-402-885/124

NIH LIBRARY
4 0146 8844

NIH Library, Building 10
National Institutes of Health
Bethesda, Md. 20205

New Book 12-19-83

DATE DUE



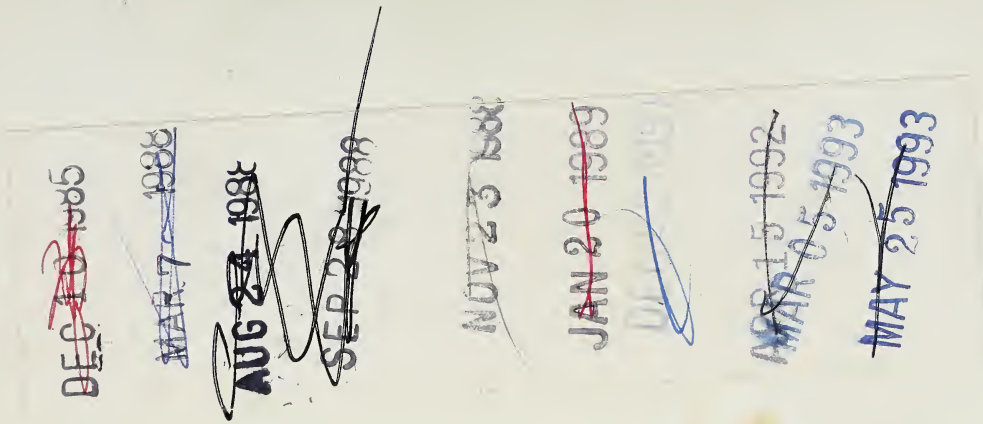
LIBRARY
NIH
Amazing Research.
Amazing Help.
http://nihlibrary.nih.gov

10 Center Drive
Bethesda, MD 20892-1150
301-496-1080

GAYLORD PRINTED IN U.S.A.

President's Commission for the Study of
Ethical Problems in Medicine and
Biomedical and Behavioral Research

Suite 555
2000 K Street, NW
Washington, DC 20006

Postage and Fees Paid
President's Commission for the Study
of Ethical Problems in Medicine and
Biomedical and Behavioral Research

Official Business
Penalty For Private Use $300